Fluorescence devices for the detection of dental caries

Macey, Richard; Walsh, Tanya; Riley, Philip; Glenny, Anne-Marie; Worthington, Helen V.; Fee, Patrick A.

Published in:
The Cochrane database of systematic reviews

DOI:
10.1002/14651858.CD013811

Publication date:
2020

Document Version
Publisher's PDF, also known as Version of record

Cochrane Library

Cochrane Database of Systematic Reviews

Fluorescence devices for the detection of dental caries (Review)

Macey R, Walsh T, Riley P, Glenny AM, Worthington HV, Fee PA, Clarkson JE, Ricketts D

Macey R, Walsh T, Riley P, Glenny A-M, Worthington HV, Fee PA, Clarkson JE, Ricketts D.
Fluorescence devices for the detection of dental caries.
Cochrane Database of Systematic Reviews 2020, Issue 12. Art. No.: CD013811.
DOI: 10.1002/14651858.CD013811.

www.cochranelibrary.com

WILEY

TABLE OF CONTENTS

[Diagnostic Test Accuracy Review]

Fluorescence devices for the detection of dental caries

Richard Macey[1], Tanya Walsh[1], Philip Riley[2], Anne-Marie Glenny[1], Helen V Worthington[2], Patrick A Fee[3], Janet E Clarkson[4], David Ricketts[5]

[1]Division of Dentistry, School of Medical Sciences, Faculty of Biology, Medicine and Health, The University of Manchester, Manchester, UK. [2]Cochrane Oral Health, Division of Dentistry, School of Medical Sciences, Faculty of Biology, Medicine and Health, The University of Manchester, Manchester, UK. [3]Dundee Dental School, University of Dundee, Dundee, UK. [4]Division of Oral Health Sciences, Dundee Dental School, University of Dundee, Dundee, UK. [5]Dundee Dental School, University of Dundee, Dundee, UK

Contact address: Tanya Walsh, tanya.walsh@manchester.ac.uk.

Editorial group: Cochrane Oral Health Group.
Publication status and date: New, published in Issue 12, 2020.

Citation: Macey R, Walsh T, Riley P, Glenny A-M, Worthington HV, Fee PA, Clarkson JE, Ricketts D. Fluorescence devices for the detection of dental caries. *Cochrane Database of Systematic Reviews* 2020, Issue 12. Art. No.: CD013811. DOI: 10.1002/14651858.CD013811.

ABSTRACT

Background

Caries is one of the most prevalent and preventable conditions worldwide. If identified early enough then non-invasive techniques can be applied, and therefore this review focusses on early caries involving the enamel surface of the tooth. The cornerstone of caries detection is a visual and tactile dental examination, however alternative methods of detection are available, and these include fluorescence-based devices. There are three categories of fluorescence-based device each primarily defined by the different wavelengths they exploit; we have labelled these groups as red, blue, and green fluorescence. These devices could support the visual examination for the detection and diagnosis of caries at an early stage of decay.

Objectives

Our primary objectives were to estimate the diagnostic test accuracy of fluorescence-based devices for the detection and diagnosis of enamel caries in children or adults. We planned to investigate the following potential sources of heterogeneity: tooth surface (occlusal, proximal, smooth surface or adjacent to a restoration); single point measurement devices versus imaging or surface assessment devices; and the prevalence of more severe disease in each study sample, at the level of caries into dentine.

Search methods

Cochrane Oral Health's Information Specialist undertook a search of the following databases: MEDLINE Ovid (1946 to 30 May 2019); Embase Ovid (1980 to 30 May 2019); US National Institutes of Health Ongoing Trials Register (ClinicalTrials.gov, to 30 May 2019); and the World Health Organization International Clinical Trials Registry Platform (to 30 May 2019). We studied reference lists as well as published systematic review articles.

Selection criteria

We included diagnostic accuracy study designs that compared a fluorescence-based device with a reference standard. This included prospective studies that evaluated the diagnostic accuracy of single index tests and studies that directly compared two or more index tests. Studies that explicitly recruited participants with caries into dentine or frank cavitation were excluded.

Data collection and analysis

Two review authors extracted data independently using a piloted study data extraction form based on the Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2). Sensitivity and specificity with 95% confidence intervals (CIs) were reported for each study. This information has been displayed as coupled forest plots and summary receiver operating characteristic (SROC) plots, displaying the

sensitivity-specificity points for each study. We estimated diagnostic accuracy using hierarchical summary receiver operating characteristic (HSROC) methods. We reported sensitivities at fixed values of specificity (median 0.78, upper quartile 0.90).

Main results

We included a total of 133 studies, 55 did not report data in the 2 x 2 format and could not be included in the meta-analysis. 79 studies which provided 114 datasets and evaluated 21,283 tooth surfaces were included in the meta-analysis. There was a high risk of bias for the participant selection domain. The index test, reference standard, and flow and timing domains all showed a high proportion of studies to be at low risk of bias. Concerns regarding the applicability of the evidence were high or unclear for all domains, the highest proportion being seen in participant selection. Selective participant recruitment, poorly defined diagnostic thresholds, and in vitro studies being non-generalisable to the clinical scenario of a routine dental examination were the main reasons for these findings. The dominance of in vitro studies also means that the information on how the results of these devices are used to support diagnosis, as opposed to pure detection, was extremely limited. There was substantial variability in the results which could not be explained by the different devices or dentition or other sources of heterogeneity that we investigated. The diagnostic odds ratio (DOR) was 14.12 (95% CI 11.17 to 17.84).

The estimated sensitivity, at a fixed median specificity of 0.78, was 0.70 (95% CI 0.64 to 0.75). In a hypothetical cohort of 1000 tooth sites or surfaces, with a prevalence of enamel caries of 57%, obtained from the included studies, the estimated sensitivity of 0.70 and specificity of 0.78 would result in 171 missed tooth sites or surfaces with enamel caries (false negatives) and 95 incorrectly classed as having early caries (false positives).

We used meta-regression to compare the accuracy of the different devices for red fluorescence (84 datasets, 14,514 tooth sites), blue fluorescence (21 datasets, 3429 tooth sites), and green fluorescence (9 datasets, 3340 tooth sites) devices. Initially, we allowed threshold, shape, and accuracy to vary according to device type by including covariates in the model. Allowing consistency of shape, removal of the covariates for accuracy had only a negligible effect ($Chi^2 = 3.91$, degrees of freedom (df) = 2, P = 0.14).

Despite the relatively large volume of evidence we rated the certainty of the evidence as low, downgraded two levels in total, for risk of bias due to limitations in the design and conduct of the included studies, indirectness arising from the high number of in vitro studies, and inconsistency due to the substantial variability of results.

Authors' conclusions

There is considerable variation in the performance of these fluorescence-based devices that could not be explained by the different wavelengths of the devices assessed, participant, or study characteristics. Blue and green fluorescence-based devices appeared to outperform red fluorescence-based devices but this difference was not supported by the results of a formal statistical comparison. The evidence base was considerable, but we were only able to include 79 studies out of 133 in the meta-analysis as estimates of sensitivity or specificity values or both could not be extracted or derived. In terms of applicability, any future studies should be carried out in a clinical setting, where difficulties of caries assessment within the oral cavity include plaque, staining, and restorations. Other considerations include the potential of fluorescence devices to be used in combination with other technologies and comparative diagnostic accuracy studies.

PLAIN LANGUAGE SUMMARY

Fluorescence devices for the detection of dental caries

Why is it important to improve dental caries (tooth decay) detection?

Dentists often aim to identify tooth decay that has already advanced to a level which needs a filling. If dentists were able to find tooth decay when it has only affected the outer layer of the tooth then it is possible to stop the decay from spreading any further and prevent the need for fillings. It is also important to avoid a false-positive result, when treatment may be provided when caries is absent.

What is the aim of this review?

This Cochrane Review aimed to find out how accurate fluorescence devices (non-invasive devices that shine a light on the surface of the tooth) are for detecting and diagnosing early tooth decay as part of the dental 'check-up' for children and adults who visit their general dentist. Researchers included 133 studies to answer this question.

What was studied in the review?

There are three different types of fluorescence device that use different types of light which we grouped as red, blue, and green fluorescence. Each device reflects more or less light depending on the amount of tooth decay, and this is measured by the device to give a score which indicates whether there is tooth decay and how severe the decay is. We studied decay on the occlusal surfaces (biting surfaces of the back teeth), the proximal surfaces (tooth surfaces that are next to each other), and the smooth surfaces.

What are the main results of the review?

The review included 133 relevant studies but 55 of these did not provide data in a format that we could use for analysis, so 79 studies with a total of 21,283 teeth were included in the analysis. Some of these studies reported on more than one type of fluorescence device, this gave us 114 sets of data. The results of these studies indicate that, in theory, if the fluorescence devices were to be used by a dentist for a routine dental examination in a group of 1000 tooth sites or surfaces, of which 574 (57%) have early tooth decay:

• an estimated 494 will have a fluorescence device result indicating tooth decay, and of these, 95 (19%) will not have tooth decay (false positive - incorrect diagnosis);
• of the 506 tooth sites or surfaces with a result indicating that tooth decay is not present, 171 (34%) will have early tooth decay (false negative - incorrect diagnosis).

Please see oralhealth.cochrane.org/fluorescence-devices-results.

We found no evidence that the devices that used different types of light (red, blue, or green fluorescence) differed in their accuracy.

How reliable are the results of the studies in this review?

We only included studies that assessed healthy teeth or those that were thought to have early tooth decay. This is because teeth with deep tooth decay would be easier to detect. However, there were some problems with how the studies were carried out. This may have resulted in the fluorescence-based devices appearing more accurate than they are. We judged the certainty of the evidence as low due to how the studies selected their participants, the large number of studies that were carried out in a laboratory setting on extracted teeth, and the variation in results reported.

Who do the results of this review apply to?

Studies included in the review were carried out in Brazil, Europe, the Middle East, Asia, North America, and Australia. A large number of studies used extracted teeth. Others were completed in dental hospitals, general dental practices, or schools. Studies were from the years 1998 and 2019.

What are the implications of this review?

Because of the wide variation in performance that cannot be easily explained the interpretation of results is difficult. The proportion of cases missed or incorrectly diagnosed as evidence of caries is relatively high. Important information was missing from many of the included studies. Any future studies should be carried out in a clinical setting, and look at the potential of fluorescence devices to be used alongside other devices.

How up-to-date is this review?

The review authors searched for and used studies published up to 30 May 2019.

SUMMARY OF FINDINGS

Summary of findings 1. Summary of findings table - main results

Question	**What is the diagnostic accuracy of fluorescence-based index tests for the detection and diagnosis of early dental caries?**				
Population	Children or adults who are presenting asymptomatically or who are suspected of having enamel caries (clinical studies); extracted teeth of children or adults (in vitro studies). Studies which intentionally included dentine and frank cavitations were excluded				
Index test	Fluorescence-based devices - including red, blue, and green fluorescence, suitable for use as an adjunct to a conventional clinical oral examination. Results of the index test were given on a continuous scale using a software algorithm				
Comparator test	Comparisons were made between fluorescence devices				
Target condition	Dental caries, at the threshold of caries in enamel				
Reference standard	Histology, enhanced visual examination with or without radiographs				
Action	Caries lesions confined to tooth enamel have the potential to be stabilised or even reversed, whereas the progression of carious lesions into the deeper aspects of dentine and pulp of the tooth will often require restorative treatment				
Diagnostic stage	Aimed at the general dental practitioner assessing regularly attending patients for early-stage caries				
Quantity of evidence	79 studies providing data for meta-analysis (133 studies included in the systematic review) (114 datasets, 21,283 tooth surfaces of which 12,138 tooth surfaces with caries at enamel threshold or greater (57% prevalence))				
Findings					
Estimated sensitivity (95% CI)[a]	0.70 (0.64 to 0.75) at median fixed specificity of 0.78; 0.60 (0.54 to 0.65) at upper quartile fixed specificity of 0.90				
DOR (95% CI)	14.12 (11.17 to 17.84)				
Effect per 1000 tooth sites or surfaces assessed	**Numbers applied to a hypothetical cohort of 1000 tooth sites or surfaces: sensitivity at fixed specificity 0.78 (95% CI)**		**Numbers applied to a hypothetical cohort of 1000 tooth sites or surfaces: sensitivity at fixed specificity 0.90 (95% CI)**		**Test accuracy Certainty of the evidence**
Outcome	**Pre-test probability 28%[b]**	**Pre-test probability 57%[b]**	**Pre-test probability 28%[b]**	**Pre-test probability 57%[b]**	
True positives (patients with early enamel caries)	196 (179 to 210)	399 (365 to 428)	168 (151 to 182)	342 (308 to 371)	⊕⊕⊝⊝ LOW

False negatives (patients incorrectly classified as not having early enamel caries)	84 (70 to 101)	171 (142 to 205)	112 (98 to 129)	228 (199 to 262)
True negatives (patients without early enamel caries)	562 (526 to 598)	335 (314 to 357)	648 (626 to 662)	387 (374 to 396)
False positives (patients incorrectly classified as having early enamel caries)	158 (122 to 194)	95 (73 to 116)	72 (58 to 94)	43 (34 to 56)

Limitations	
Risk of bias	Of the 79 studies included in the meta-analysis: patient selection was registered as having a low risk of bias due to the use of consecutive or random sampling in 9 studies, avoiding a case-control design (79 studies), and avoiding inappropriate exclusions (64 studies). A low risk of bias was observed when the index tests could not be influenced by the reference standard (61 studies) and where thresholds were clearly reported (50 studies). There was a low risk of bias when the reference standard correctly classified the target condition (49 studies) and where the reference standard was interpreted without knowledge of the index test (49 studies). Low risk of bias was allocated for flow and timing when there was no concern regarding the interval between tests (79 studies), the same reference standard was used for all tooth surfaces (68 studies), and all tooth surfaces were reported in the analysis (65 studies) Risk of bias for all results included in the review (133) is reported in the main text
Applicability of evidence to the review question	Patient selection was considered to be a high concern in studies where extracted teeth were used (78 studies). Inappropriately defined thresholds for the index test also resulted in high concern for applicability, this occurred when early enamel caries were categorised with the sound teeth (1 study) and for reference standard (4 studies). The dominance of in vitro studies also means that information on how the results of these devices are used to support diagnosis, as opposed to pure detection, is extremely limited
Certainty of the evidence	We downgraded the certainty of the evidence by 2 levels in total for risk of bias due to limitations in the design and conduct of the included studies, indirectness arising from the high number of in vitro studies, and inconsistency due to the substantial variability in results

[a]2 illustrative examples of points on the SROC curve fixed at the median specificity of 0.78 followed by upper quartile specificity of 0.90.
[b]Hypothetical cohorts of 1000 lesions are presented for numbers estimated at prevalence of 28% and 57% of enamel caries prevalence. Based on consultation with clinical colleagues, the lower prevalence figure addresses the concern that the higher prevalences of 57% are not representative of the general population and is taken from the level of cavitated teeth in the UK Adult Dental Health Survey (Steele 2011). The higher prevalence figure is taken from the total number of observed caries in the included studies divided by the total number of included tooth surfaces.
CI: confidence interval; **DOR:** diagnostic odds ratio; **SROC:** summary receiver operating characteristic plot.

BACKGROUND

Cochrane Oral Health (COH) has undertaken several systematic reviews of diagnostic test accuracy (DTA) on the detection and diagnosis of dental caries. The suite of systematic reviews forms part of a UK National Institute for Health Research (NIHR) Cochrane Programme Grant Scheme and involved collaboration with the Complex Reviews Support Unit. The reviews follow standard Cochrane DTA methodology and have been differentiated according to the index test under evaluation. A generic protocol served as the basis for the suite of systematic reviews (Macey 2018).

Caries is an entire disease process, which can be stabilised and sometimes reversed if diagnosed and treated early on in the disease process (Fejerskov 2015; Pitts 2009). Most high-income countries around the world have evidenced a reduction in caries incidence in children and adolescents, and in some Scandinavian countries prevention programmes have almost eradicated caries, but such activities have not been widely replicated in other locations (Pitts 2017). Despite this reduction, the 2015 Global Burden of Disease study identified dental caries as the most prevalent, preventable condition worldwide (Feigin 2016; Kassebaum 2015), affecting 60% to 90% of children and the majority of adults of the world's population (Dye 2015; Petersen 2005). Furthermore, despite a reduction in caries in many industrialised countries, the global incidence of untreated caries was reported to be 2.4 billion in 2010 (Feigin 2016; Kassebaum 2015; World Health Organization 2017) and continues to increase year on year. In the UK, the primary reason for childhood (aged 5 years to 9 years) hospital admissions is for the extraction of teeth (Public Health England 2014). Longitudinal studies have shown that those who experience caries early in childhood will have an increased risk of severe caries in later life, and that the disease trajectory will be steeper than those without early caries experience (Broadbent 2008; Hall-Scullin 2017).

Untreated caries can lead to episodes of severe pain and infection, often requiring treatment with antibiotics. Dental anxiety resulting from untreated caries and the subsequent need for more invasive management, can adversely affect a person's future willingness to visit their dentist, leading to a downward spiral of oral disease (Milsom 2003; Thomson 2000). If left to progress, treatment options are limited to restoration or extraction, requiring repeated visits to a dental surgery or even to a hospital (Featherstone 2004; Fejerskov 2015; Kidd 2004).

The cost of treating caries is high. In the UK alone, the National Health Service (NHS) spends around GBP half a billion every year in treating the disease. Hidden costs also exist, and the related productivity losses are high, estimated at USD 27 billion globally in 2010 (Listl 2015).

Caries detection and diagnosis will usually be undertaken at a routine dental examination, by a general dental practitioner, in patients who are presenting asymptomatically. However, caries detection can additionally be employed in secondary care settings, school or community screening projects, and epidemiology or research studies (Braga 2009; Jones 2017). The traditional method of detecting dental caries in clinical practice is a visual-tactile examination often with supporting radiographic investigations. This combination of methods is believed to be successful at detecting caries that has progressed into dentine and reached a threshold where a restoration may be necessary (Kidd 2004). However, the detection of caries earlier in the disease continuum could lead to stabilisation of disease or even possible remineralisation of the tooth surface, thus preventing the patient from entering a lifelong cycle of restoration (Pitts 2017), but early caries is difficult to detect visually, and the use of radiographs provides only limited ability to detect small changes in dental enamel (Ismail 2007).

Detection and diagnosis at the initial (non-cavitated) and moderate levels of caries is fundamental in achieving the promotion of oral health and prevention of oral disease (Fejerskov 2015; Ismail 2013). The prevalence of this early caries state is not often reported in dental epidemiology, most reports preferring to focus on cavitated/dentinal lesions which may be easier to detect, for example, the most recent UK Adult Dental Health survey reported 31% of the sample having untreated caries into dentine (Steele 2011; White 2012), and a US study reported levels of cavities at 15.3% in 12- to 19-year olds (Dye 2015). However, one UK survey of children identified "clinical decay experience" which incorporated any enamel breakdown and all other forms of caries and reported a prevalence of 63% in 15-year olds (Children's Dental Health Survey 2013).

A wide variety of management options are available under NHS care at these different thresholds of disease, ranging from non-operative preventive strategies such as improved oral hygiene, reduced sugar diet and application of topical fluoride to minimally invasive treatments (e.g. sealing the affected surface of the tooth, or 'infiltrating' the demineralised tissue with resins) for initial caries, through to selective caries removal and restoration for extensive lesions. With advances in technology over the last two decades, additional methods of detection have become available, such as advancements in radiography and the development of fluorescence, transillumination, and electrical conductance devices. These could potentially aid the detection and diagnosis of caries at an early stage of decay. This would afford the patient the opportunity of a less invasive treatment with less destruction of tooth tissue and potentially result in a reduced cost of care to the patient and healthcare services.

Target condition being diagnosed

The term dental caries is used to describe the mechanism which can ultimately lead to the breakdown of the tooth surface which results from an imbalance in the activity within the biofilm (or dental plaque) on the surface of the tooth within the oral cavity (Kidd 2016). This imbalance is due to bacterial breakdown of sugars in the diet which leads to the production of acid and subsequent demineralisation of the tooth. Disease progression can be moderated by improved oral hygiene practices together with the influx of fluoride from toothpaste and other available fluoride sources. However, the levels of sugar consumption observed in many populations will often outweigh the benefits of fluoride (Hse 2015). Ultimately, carious lesions may develop and destroy the structure of the tooth.

The most common surfaces for caries to manifest are on the occlusal (biting) surfaces or the proximal surfaces (tooth surface which face an adjacent tooth); although smooth surfaces on the flat exterior of teeth adjacent to the tongue, cheeks, and lips can be affected. The severity of the disease is defined by the depth of demineralisation of the tooth's structure and whether the lesion is active or arrested. Caries presenting at levels into tooth enamel can

potentially be stabilised or even reversed, whereas the progression of carious lesions into the dentine and pulp of the tooth will often require restoration (Bakhshandeh 2018; Kidd 2004).

Assessment of disease severity traditionally used in epidemiological and research studies has historically employed some variant of the DMFT (decayed, missing, and filled teeth) scale. Within the D (decayed) component there are four clinically detectable thresholds applied as indicators for diagnosis and treatment planning, often labelled as D_1, D_2, D_3, and D_4 (Anaise 1984) (Additional Table 1). Typically the D_3 threshold, with only lesions extending into dentine classed as carious, has been used to determine the presence of caries (Pitts 1988; Shoaib 2009). These four categories have formed the basis for expanded caries indices based on visual characteristics such as the International Caries Detection and Assessment System (ICDAS) (Ekstrand 2007; Ismail 2007). Other available systems include: the Nyvad system (Nyvad 1999); Ekstrand-Ricketts-Kidd (ERK) system (Ekstrand 1997); British Association for the Study of Community Dentistry (BASCD) (Pitts 1997); the Dundee Selectable Threshold Method for caries diagnosis (DSTM) (Fyffe 2000); and the American Dental Association Caries Classification System for clinical practice (Young 2015). The ICDAS and DSTM systems both provide the opportunity to investigate initial caries (into enamel) which may confer benefits for preventative or non-operative treatment.

Treatment of caries

There are many varied treatment options available to the dental clinician, dependent on the thresholds of observed disease. Initial caries can be managed without surgical intervention using approaches such as plaque control, dietary advice, and application of fluoride to remineralise the tooth surface and prevent further progression (Kidd 2016). Minimally invasive treatments for initial caries are available, such as sealing the affected surface of the tooth, or 'infiltrating' the demineralised tissue with resins. High-risk patients with severe caries may require selective caries removal and restoration of extensive lesions.

A caries management pathway, informed by diagnostic information, can be beneficial in guiding the clinician towards prevention or a treatment plan. One recently developed care pathway is the International Caries Classification and Management System (ICCMS) (Ismail 2015). The system presents three forms of management in the care pathway:

- when dentition is sound the clinician proceeds with preventative strategies to prevent sound surfaces from developing caries;
- non-invasive treatment of the lesion to arrest the decay process and encourage remineralisation, preventing initial lesions from progressing to cavitated decay; and
- management of more severe caries through excavation and restoration or potentially extraction.

At the core of this care pathway is the ability to detect early caries accurately and optimise the preventative strategies through tooth tissue-preserving excavation methods, and restoration or potentially extraction in more severe cases. The detection and diagnosis of early caries remain challenging, and the likelihood of undiagnosed early disease is high (Ekstrand 1997). In such instances, the opportunity for preventing initial lesions from progressing to cavitated decay, or even reversing the disease process, is missed, and disease progresses to cavitated decay where restoration is required (Ekstrand 1998).

Index test(s)

The cornerstone of caries detection is a visual and tactile dental examination, and the ability of clinicians to accurately detect disease in this way has been researched for over half a century (Backer Dirks 1951). Many devices for the detection and diagnosis now exist and may be suitable at different stages of the care pathway (Bloemendal 2004; Fyffe 2000). This review investigates fluorescence-based devices that aim to measure the mineral content of the tooth according to changing fluorescence identified using light with various wavelengths according to the device used (e.g. 405 nm for quantitative light-induced fluorescence (QLF) and 655 nm for DIAGNOdent) (Kim 2019; Neuhaus 2019). Macey 2018 provide details of the other index tests being investigated in this series of systematic reviews.

We included three categories of fluorescence index test each primarily defined by the different wavelengths exploited by the devices.

- Red fluorescence: these devices use a small laser with an excitation wavelength greater than 655 nm. The tip of the device emits the excitation light and collects the resultant fluorescence and works on the principle that carious tissue creates more emitted fluorescence than sound tissue through the fluorescence of bacterial by-products (porphyrins) (Pretty 2006). These devices include: DIAGNOdent and DIAGNOdent pen (KaVo, Biberach, Germany) that feedback results via the device's display on a continuous scale (minimum 1 to maximum 99); MidWest (DENTSPLY Professional, New York, USA) emits sound and light (green/red) if caries is detected; and the Canary System (Quantum Dental Technologies Inc, Toronto, Ontario, Canada) which displays a number on a scale from 0 to 100 where 0 to 20 is deemed to be healthy (Amaechi 2019; Lussi 1999; Lussi 2001; Neuhaus 2019; Rodrigues 2011).
- Blue fluorescence: these devices operate at wavelengths between 400 nm and 450 nm at the blue/violet end of the visible light spectrum and create luminescence in regions where there is bacterial activity which is often indicative of dental caries; while the sound or healthy areas of the tooth continue to fluoresce green (Rodrigues 2011). The devices in this group rely on bespoke software to provide an image of the luminescence regions, examples are VistaProof (Durr Dental, Bietigheim-Bissingen, Germany), SoproLife (ACTEON Group, La Ciotat, France), and Spectra (Air Techniques, Melville, New York, USA) which use bespoke software packages to produce a digital image of the tooth which is interpreted by the operator. The devices use different wavelengths of light (405 nm versus 450 nm) however their mode of action is similar. VistaProof uses software to create a numeric score between 0 and 5 (Achilleos 2013), SoproLife relies on the operator interpreting the findings of the imaging program and allocating to one of six groups that range from sound to visible dentine (Rechmann 2012), Spectra provides a numeric and colour category ranging from sound to dentine lesions (Graye 2012).
- Green fluorescence: includes devices that use QLF, these rely on the characteristics of fluorescence at the green-yellow end of the spectrum (370 nm) (Angmar-Månsson 2001). This is emitted or refracted to the device and a measurement is taken,

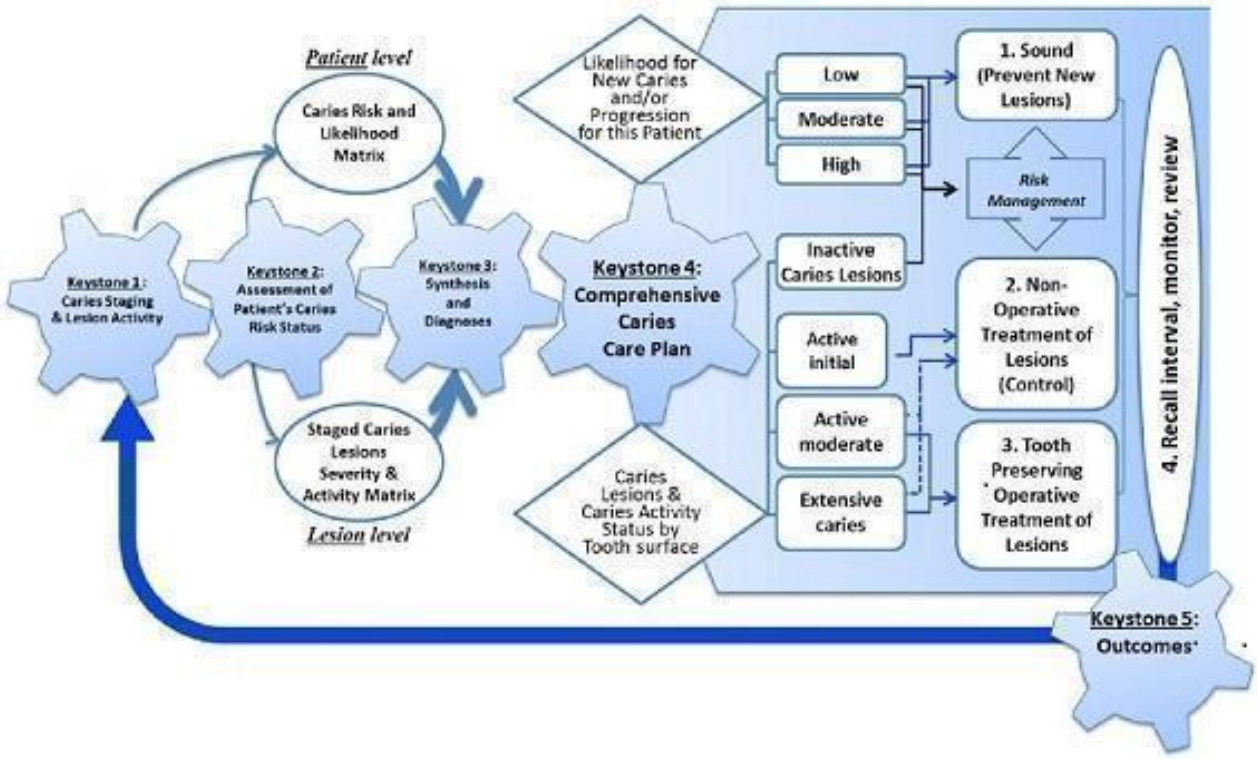
which by definition is the tooth's "quantitative light-induced fluorescence" and can be measured in terms of an average loss of fluorescence denoting lesion depth (often labelled ΔF and allocated to a point on a numeric scale) (Kim 2019; Neuhaus 2016).

Clinical pathway

The process from a dental patient attending for a routine examination and a caries assessment being undertaken potentially has four intertwined phases: screening, detection, diagnosis, and treatment planning. If the presenting patient is at some risk of disease but seemingly asymptomatic then this can be considered as a screening exercise (Wilson 1968) to detect initial caries in individuals who do not yet have symptoms. Since caries is a dynamic process the pure detection of the disease at a single time point is not sufficient to inform the future care of the patient, and additionally the depth and severity of demineralisation, allied to a decision on the caries activity levels, must be combined to reach a diagnosis (Ismail 2004; Nyvad 1997). This diagnosis then feeds into a caries management pathway once the patient's history, personal oral care, and risk factors have been considered. A comprehensive methodology has been developed, the International Caries Classification and Management System (ICCMS™), that "helps practitioners to intuitively and systematically collect and analyze personal and clinical data to develop comprehensive patient care plans" (Ismail 2015) that go beyond restorative care.

Figure 1 presents the key elements of the ICCMS. This Cochrane Review aims to inform the process at 'Keystone 3' where diagnosis is an indefinable component.

Figure 1. Keystones of the International Caries Classification and Management System (ICCMS™). Copyright© 2018 Ismail AI, Pitts NB, Tellez M. The International Caries Classification and Management System (ICCMS™) an example of a caries management pathway. *BMC Oral Health* 2015;15(Suppl 1):S9. Reproduced with permission.



Role of index test(s)

The role of the proposed fluorescent devices may vary according to whether the purpose of the examination is detection or diagnosis. For detection or case-finding, the fluorescence-based device could, in theory, be used as a standalone test. However, some form of implicit visual assessment will be required for correct placement of the device. This is particularly so for 'point-based' devices which have a relatively narrow area of focus. In clinical practice, a conventional oral examination would always be undertaken as part of the clinical examination, and as such, it is unlikely that any of the index tests under evaluation would be used as a complete replacement for the combined activities of detection and diagnosis of initial decay. Supplementing the visual-tactile examination with an index test could support the detection of initial decay. The index tests could also have a triage role in assisting the general dentist to more accurately assess signs of uncertain clinical significance. The information from caries detection (including assessment of the severity of disease) will be an integral part of a person's diagnosis, which additionally incorporates their clinical history, risk factors, and treatment planning protocols.

Alternative test(s)

Alternative tests include.

- Comprehensive visual or visual-tactile examination with a detailed classification system: identifying caries according to visual appearance, aided by a dental mirror and sometimes a probe, on clean and dry teeth.
- Radiography: bitewing radiology is most commonly used. Other techniques include: subtraction radiography which produces a semi-automated method for monitoring progression of lesions (Ellwood 1997; Wenzel 2006) and cone-beam computed technology (CBCT) which provides a three dimensional image which appears to offer great potential for diagnosis with increased levels of radiation (Horner 2009).
- Fibre-optic transillumination (FOTI) which uses a light emitted from a handheld device that when placed directly onto the tooth illuminates the tooth (Pretty 2006). Any demineralisation should appear as shadows in the tooth due to the disruption of the tooth's structure due to caries.
- Electrical conductance: the demineralisation of the tooth is reported to affect the tooth's electrical conductance. This is measured by placing a probe on the tooth which measures any potentially higher conductivity which occurs due to carious lesions being filled with saliva (Tam 2001).

For more details please see the generic protocol for this review (Macey 2018).

Rationale

Despite technological advancement, caries detection is typically based upon information from a visual-tactile clinical examination with or without radiographs. Bader 2002 completed an extensive literature review of in vitro caries detection studies investigating visual, dental imaging, fibre-optic, electrical conductance, and fluorescence in primary and permanent dentition. The review was restricted to studies that included a histological reference standard and grouped studies according to index test, disease threshold (enamel or dentinal lesions), and tooth surface (occlusal or proximal); no meta-analysis was undertaken, and the authors graded the quality of the available evidence as low (Bader 2002). Two years later the same authors published a review focusing on fluorescence devices. Despite an increase in the number of eligible studies in the intervening years, the authors determined that it was still not possible to carry out a meta-analysis and raised concerns over the propensity of the fluorescence devices for decreasing specificity as sensitivity improved (Bader 2004). These two reviews predate the development of meta-analysis methods for DTA reviews recommended in the *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy* (Deeks 2013). A subsequent systematic review investigated the accuracy of fluorescence devices, and included studies of the primary and permanent dentition, occlusal and proximal surfaces, with reference standards of histology, operative, visual examination, and dental imaging (Gimenez 2013). We aimed to build upon existing research in caries detection and diagnosis by expanding the search strategy to capture all relevant evidence, applying appropriate hierarchical meta-analytical models (Dinnes 2016), and assessing the body of evidence using GRADE (Schünemann 2020; Schünemann 2020a) to facilitate the production of 'Summary of findings' tables.

OBJECTIVES

Primary objectives

To determine the diagnostic accuracy of fluorescence-based index tests used alone or in combination with other tests for the detection and diagnosis of coronal dental caries in children and adults. We aimed to evaluate the comparative accuracy of red, blue, and green fluorescence-based devices; these included DIAGNOdent, DIAGNOdent pen, SoproLife, VistaProof, and quantitative light-induced fluorescence (QLF). The specific research questions addressed in this systematic review were.

- What is the diagnostic test accuracy of fluorescence-based tests for detection or diagnosis in different populations (children: primary/mixed dentition, adolescents: immature permanent dentition, or adults: mature permanent dentition), and when tested against different reference standards.
    - What is the diagnostic test accuracy of each of the three groups of fluorescence-based index tests compared to an appropriate reference standard for detecting and diagnosing initial stage decay on the occlusal, proximal, and smooth tooth surfaces?
    - Do measures of sensitivity and specificity for single tests differ from the sensitivity and specificity of tests used in combination (fluorescence test either individually or combined with a visual examination)? Is there a benefit to using more than one index test as opposed to a single test?

Secondary objectives

We aimed to investigate the following potential sources of heterogeneity.

- Recruited population - children: primary/mixed dentition, adolescents: immature permanent dentition, or adults: mature permanent dentition.
- Prevalence of caries into dentine in the study sample.
- Tooth surface being reported (occlusal, proximal, smooth surface or adjacent to a restoration).
- Reference standards - in vitro studies commonly use histology as the reference standard.
- Consideration of point measurement devices versus imaging or surface assessment devices.

METHODS

Criteria for considering studies for this review

Types of studies

We considered diagnostic accuracy study designs that were:

- studies with a single set of inclusion criteria that compared a fluorescence diagnostic test with a reference standard. We included prospective studies that evaluated the diagnostic accuracy of single index tests, and studies that directly compared two or more index tests;
- randomised controlled trials (RCTs) of the diagnostic test accuracy of one or more index tests in comparison, or versus a no test option;

- 'case-control' type accuracy studies where different sets of criteria were used to recruit those with or without the target condition, although prone to bias some innovative tests may be identifiable through this design only and this eligibility criterion may provide an opportunity to report them, these studies would not be included in the primary analysis;
- studies reporting at both the patient and tooth or tooth surface level were included, however only those reporting at the tooth surface level would be included in the primary analysis.

In vitro and in vivo studies were eligible for inclusion. In vitro studies use teeth that have been extracted prior to the start of the study. The index test is carried out on extracted teeth, albeit in a scenario which is not representative of the typical clinical setting, and will typically be followed by a reference standard of histology. In vivo studies recruit participants and conduct index tests with the teeth in the oral cavity. The reference standard is usually enhanced clinical examination or excavation. In some cases the reference standard is histology, for example when a study has been conducted with participants who have teeth indicated for extraction due to orthodontic or third molar indications, periodontal diseases, or children with teeth that are due to exfoliate naturally.

We excluded studies where:

- artificially created carious lesions were used in the testing procedure;
- an index test was used during the excavation of dental caries to ascertain the optimum depth of excavation.

Participants

Participants who are seemingly asymptomatic for dental caries, including those who may have carious lesions that are undetected at the point of enrolment. Studies that explicitly recruited participants with caries into dentine or frank cavitation were excluded. We also excluded studies where participants were referred to secondary care for restorative treatment, as there is a likelihood that advanced caries (into dentine or pulp) would be present and readily detectable without the need for the index tests investigated in this review.

Studies recruiting children, adolescents, and adults were all eligible for inclusion. This allowed for the analysis of the diagnostic test accuracy of index tests for primary, mixed, and permanent dentition.

Index tests

Fluorescence-based devices: incorporating a variety of devices that included laser-based detection. Devices may have been used as an adjunct to a conventional visual examination and require an operator judgement or generate a conclusion via a software algorithm. There was considerable variation in the positivity thresholds used across the different fluorescence-based devices. The devices that provided a numeric output on a continuous scale were often interpreted at different thresholds, but where multiple thresholds were reported within a study report we extracted data at the pre-specified manufacturers' threshold wherever possible.

These index tests were completed on intact teeth and could be used as an adjunct or replacement for aspects of the current examination. The intention was to assess the index tests in isolation wherever possible, otherwise the result of one index test may have influenced another. Where multiple index tests were used as a combined index test these studies were reported separately.

Where studies used multiple examiners we extracted the results for the most appropriate examiner to the research question. For example, if the study used dental students, general dental practitioners, and restorative consultants, then the results of the general dental practitioners were extracted. In the scenario where multiple examiners showed similar skills and experience then the mean sensitivity and specificity results were extracted. If this was not available then the reported results from the first examiner were extracted.

Studies that investigated a standard clinical oral examination with an adjunct of fluorescence were included if the diagnostic information relating to fluorescence could be isolated from the other test. If the study reported a combined interpretation of both methods and if the review included sufficient numbers of combined tests, then we planned to create a subgroup of these combined tests.

Target conditions

Coronal caries: initial stage decay, defined as initial or incipient caries or non-cavitated lesions. Specifically where there is a detectable change in enamel evident which is not thought to have progressed into dentine on occlusal, proximal surfaces, and smooth surfaces.

Reference standards

Several different reference standards have been used in primary diagnostic test accuracy (DTA) studies for dental caries. The only way of achieving a true diagnosis of caries presence and severity is to extract and section the tooth and perform a histological assessment (Downer 1975; Kidd 2004). This would not be ethically reasonable to undertake on a healthy population in clinical (in vivo) studies, but is acceptable and widely used in in vitro studies conducted on previously extracted teeth. The only scenario where histology can be a viable scenario for clinical studies undertaken in a primary or secondary care setting would be where a tooth has been identified as requiring extraction (ideally for a non-caries related reason such as orthodontic or third molar extraction), and the index test could be applied before the extraction, followed by the reference standard of histology. However, this would bring into question the study's broader external validity as these types of studies are most likely to occur in adolescents or young adults and who are therefore not representative of the wider population.

Alternatives to extraction and histological assessment are operative exploration, where a clinician removes caries with a dental burr (drill) in preparation for restoration and reports the depth of decay. This technique would be acceptable as a reference standard for patients with caries of severity where restoration is required, but would not be ethical for caries-free patients or those with early caries since non-restorative treatment could be provided. A different reference standard would be required for these early lesions, the possibilities available are limited to an enhanced visual examination or radiographic tests. Studies that only used an enhanced visual or radiographic examination were included in the review as they have the benefit of allowing studies to be conducted in a clinical setting, however, their limitations in providing a true classification of disease would be identified

in the quality appraisal. Some primary studies have employed a composite reference standard based on the results of information from multiple sources.

A period of up to three months between the index test and the reference standard was deemed acceptable.

Search methods for identification of studies

Electronic searches

Cochrane Oral Health's Information Specialist conducted systematic searches in the following databases without language or publication status restrictions:

- MEDLINE Ovid (1946 to 30 May 2019) (Appendix 1);
- Embase Ovid (1980 to 30 May 2019) (Appendix 2).

Searching other resources

The following trial registries were searched for ongoing studies:

- US National Institutes of Health Ongoing Trials Register ClinicalTrials.gov (clinicaltrials.gov; searched 30 May 2019) (Appendix 3);
- World Health Organization International Clinical Trials Registry Platform (apps.who.int/trialsearch; searched 30 May 2019) (Appendix 4).

We searched the reference lists of included papers and previously published systematic reviews for additional publications not identified in the electronic searches.

We checked that none of the included studies had been retracted due to error or fraud.

Data collection and analysis

Selection of studies

Two review authors independently screened and assessed the results of all searches for inclusion. Any disagreements were resolved through discussion and, where necessary, consultation with another clinical or methodological member of the author team. Studies that met the criteria but that did not report the data in the format of a 2 x 2 contingency table were still included. In such instances, the study authors were contacted and the required data requested. An adapted PRISMA flowchart was used to report the study selection process (McInnes 2018).

Data extraction and management

Two review authors independently extracted data. A piloted study data extraction form based on the review inclusion criteria was developed and applied to 10 eligible studies. Disagreements were resolved through discussion with other members of the review team. Where data were reported for both occlusal and proximal surfaces the data were extracted separately for the different surfaces. Study authors were contacted to obtain missing data or characteristics which were not evident in the published paper.

We recorded the following data for each study:

- sample characteristics (age, sex, socioeconomic status, risk factors where stated, number of patients/carious lesions, lesion location, disease prevalence - at enamel and dentine thresholds);
- study setting (country, type of facility);
- the type of index test(s) used (category (i.e. red, blue, or green fluorescence), the device used, mode of action, conditions (i.e. clean/dried teeth), positivity threshold);
- study information (design, reference standard, case definition, training and calibration of personnel);
- study results (true positive, true negative, false positive, false negative, any equivocal results).

Assessment of methodological quality

We used the Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2) to assess the risk of bias and applicability of the eligible primary studies over the four domains of participant selection, index test, reference standard, and flow and timing (Whiting 2011), tailored for this review. 'Review specific' descriptions of how the QUADAS-2 items were contextualised and implemented are detailed in the accompanying checklist (Additional Table 2).

A 'Risk of bias' judgement ('high', 'low' or 'unclear') was made for each domain for each study. Broadly, if the answers to all signalling questions within a domain were judged as 'yes' (indicating low risk of bias for each question) then the domain was judged to be at low risk of bias. If any signalling question was judged as 'no', indicating a high risk of bias, the domain was scored as high risk of bias. Concerns regarding applicability were then completed for the participant selection, index test, and reference standard domains. There was some flexibility within this assessment framework which developed during the data extraction process and is detailed below.

Participant selection domain (1)

The selection of patients has a fundamental effect on the ability of an index test to detect caries. The disease categories of sound and enamel caries needed be represented in the sample and the age range of patients needed to be reported to form a complete appraisal of the index test's potential to correctly classify disease in different populations.

It was acceptable for studies to focus on a particular surface (occlusal/proximal) or age group (children/adults). Given that the primary objective centred on early enamel lesions studies should be reporting on this stage of the disease process. It was vital that within the chosen population all participants or teeth meeting the eligibility criteria should be provided with an equal or random opportunity to be included. Inappropriate exclusion may lead to an over or underestimation of the test's ability to detect disease, thus affecting the internal validity of the study.

All studies should have fully reported the methods used to select teeth. Ideally, a random or consecutive selection would be used and the procedure explicitly reported. Additionally, the prevalence of the different levels of disease severity should be reported. This information was used to inform the applicability of this test to a wider population.

Study results should be reported at the tooth or surface level, as apposed to the patient level, which has the potential for the index test and reference standard to be report on different sites within the same mouth.

Index test domain (2)

The nature of the fluorescence index tests and the visual presentation of the disease means that it should be feasible to ensure that the index test is conducted prior to the reference standard. Logically, the fluorescence tests had to be completed before the extraction of a tooth for any histological analysis, or before in situ excavation of a tooth is undertaken. This order of presentation (index test followed by reference standard) ensured that the index test was not influenced by the results of the reference standard. The fluorescence-based index tests generally used a device which reported a numerical value on a continuous scale. Where multiple index tests were used and where the fluorescence-based test was conducted after other index test(s) (e.g. radiograph), the objective reading and reporting of the fluorescence-based device mean that the results would not be influenced by preceding tests.

The threshold of disease positive and negative should be presented before any analysis, ideally by using the manufacturer's recommended settings or thresholds recommended by previously validated studies. Studies may have been designed to calculate the optimum threshold for a device but this will introduce bias. It is unlikely that studies will have utilised multiple index test examiners for the assessment of different disease severity or where they have it is probable that they each score all of the thresholds and are included for validation of the test. However, the inclusion of a signalling question here allowed for the identification of studies that have achieved this and provided data to inform future discussions.

Reference standard domain (3)

If the reference standard was an enhanced visual examination or radiograph then it should be completed by an examiner different to the index test, as the subjectivity of this type of reference standard could be compromised by knowledge of the index test results. An exception was built in for this signalling question because where the tooth has been extracted, sectioned and prepared for histological evaluation it is extremely unlikely that the examiner would be able to recall the specific tooth or participant and the results from the index test results. Time delays between index test and reference standard should be under three months for in vivo studies.

Ideally, each participant within a study would have received the same reference test. This is possible in an in vitro setting as a histological assessment can be applied to each selected, extracted tooth. In vivo studies may have applied the same reference standard by using enhanced visual examination or radiograph to all participants. If a study allocated participants or specific teeth to different reference standards then the reasons for this differential allocation should have been explicitly reported. All reference standards should have been completed without knowledge of the index test results.

Flow and timing domain (4)

The index test should be conducted before the reference standard. If the reference standard used is enhanced visual, radiograph, or excavation then there should be less than three months between index test and reference standard. Caries is a slow-growing disease so minimal changes should be experienced within this time frame. All observations should receive both an index test and reference standard. There are studies which report some teeth having an index test but not a reference standard; if a reason is clearly reported, such as teeth being broken during sectioning, then this would not influence the risk of bias decision.

Statistical analysis and data synthesis

The threshold of interest was between sound teeth and initial/early/enamel caries. This effectively created two groups, a positive group with any caries from early to advanced and a negative group of sound or healthy teeth. Estimates of diagnostic accuracy were expressed as sensitivity and specificity with 95% confidence intervals for each study and each available data point if the study reported multiple index tests, dentition (primary/permanent) or tooth surfaces (occlusal/proximal/smooth). We displayed this information as coupled forest plots and summary receiver operating characteristic (SROC) plots. When there were two or more test results reported in the same study, we included them as separate datasets, since the unit of analysis was the test result, not the patient.

Hierarchical models were used for data synthesis. The data were extracted for the target condition of early caries (caries into enamel). This target condition has been consistently used across the series of DTA caries reviews. A meta-analysis was conducted to combine the results of studies for each index test using the hierarchical summary ROC (HSROC) approach to estimate the expected values of sensitivity and specificity (Macaskill 2010). A summary curve using the HSROC model (Rutter 2001) was used to summarise the results since the devices provided a numeric output on a continuous scale and often interpreted these at different cut-offs. Consequently, it was not possible to apply a common threshold for analysis. An HSROC model was used to estimate a summary curve with parameter estimates for threshold, shape and accuracy, for all available datasets with no restrictions on dentition, tooth surface, reference standard, or prevalence of caries into dentine (D_3).

It was not possible to produce estimates of sensitivity and specificity as summary operating points with confidence and prediction regions on SROC plots with 95% confidence regions since the output of the HSROC model is the summary ROC curve. In the absence of clinical consensus of key values of specificity, we summarised the analysis using the median and upper quartile reported specificity and the corresponding estimate of sensitivity, along with the diagnostic odds ratio (DOR) with 95% confidence intervals (Takwoingi 2015). To allow for the analysis of false positives and false negatives we computed the sensitivity at the point on the SROC curve with fixed values of specificity of 0.78 and 0.90 (the median and upper quartile values from of all included datasets). These results are only included as examples of potential sensitivity and specificity pairings and should not be reported or interpreted formally as the summary points.

We made comparisons between the three device categories (blue, green, and red fluorescence) by comparing summary ROC curves (Takwoingi 2010). Initially, we allowed threshold, shape, and accuracy to vary according to device type by including covariates in the model (most complex model). Differences in the shapes of the summary curves were explored by removing the covariates for shape and comparing the results of this model to those of the complex model. Parameter estimates for the model assuming a common or different shape were used to generate HSROC curves

for the three categories as appropriate. If the different devices were observed to have a common shape then the model was further simplified by removing the covariates for accuracy, to determine whether the accuracy of the different devices differed in comparison with the previous model. The likelihood ratio test was applied to formally assess the significance of any model comparisons (Macaskill 2010).

The numbers generated for a hypothetical cohort of 1000 tooth sites or surfaces are reported in the 'Summary of findings' table along with the corresponding true positives, false negatives, false positives, and true negatives. The higher prevalence value was taken from the total number of enamel lesions in the included studies divided by the total number of included tooth surfaces. The lower prevalence figure was taken from the UK Adult Dental Health Survey (Steele 2011) and was used to address clinical considerations that the higher prevalence value of enamel caries reported in the primary studies, particularly in the in vitro studies, were not representative of that observed in the general population.

We used Review Manager 5 (Review Manager 2020), the NLMIXED procedure and the MetaDAS macro (Takwoingi 2010) in SAS 9.4 for Windows to carry out the analyses.

Investigations of heterogeneity

We initially inspected the clinical and methodological characteristics of the included studies, coupled forest plots, and summary ROC plots to form the basis of the assessment of heterogeneity. Where sufficient numbers of studies allowed, meta-regression analyses were undertaken to explore possible sources of heterogeneity. Formal model comparisons were compared using a likelihood ratio test to determine the statistical significance of adding each potential source of heterogeneity (covariate) to the HSROC model. Model comparisons proceeded as for the comparison of different tests above i.e. fit a complex model allowing shape, threshold, and accuracy to differ according to the source of heterogeneity, and assess the impact of the removal of the covariates for shape. If a common shape can be assumed then explore the impact of the removal of the covariates for accuracy. Each potential source of heterogeneity was analysed separately.

All investigations of heterogeneity were reported to aid interpretation of the results.

The sources of heterogeneity included (specified a priori).

Population

- Children or adults; the detection of disease in the different dentition of children or adolescents will affect the stage at which the disease is identified and treatment options which would be considered.
- Tooth surface being evaluated (occlusal, proximal, smooth surface or adjacent to a restoration).
- Prevalence of caries into dentine in each study sample.

Index test

- Consideration of point measurement devices versus imaging or surface assessment devices.

Reference standard

- Reference standard used: histology, excavation, enhanced visual examination, or radiograph.

Sensitivity analyses

Where a sufficient number of studies investigated the same index test, we assessed the impact of study quality on the sensitivity and specificity results.

Assessment of reporting bias

Methods currently available to assess reporting or publication bias for diagnostic studies may lead to uncertainty and misleading results from funnel plots (Deeks 2005; Leeflang 2008), therefore we did not carry out any tests of reporting bias.

Presentation of main results

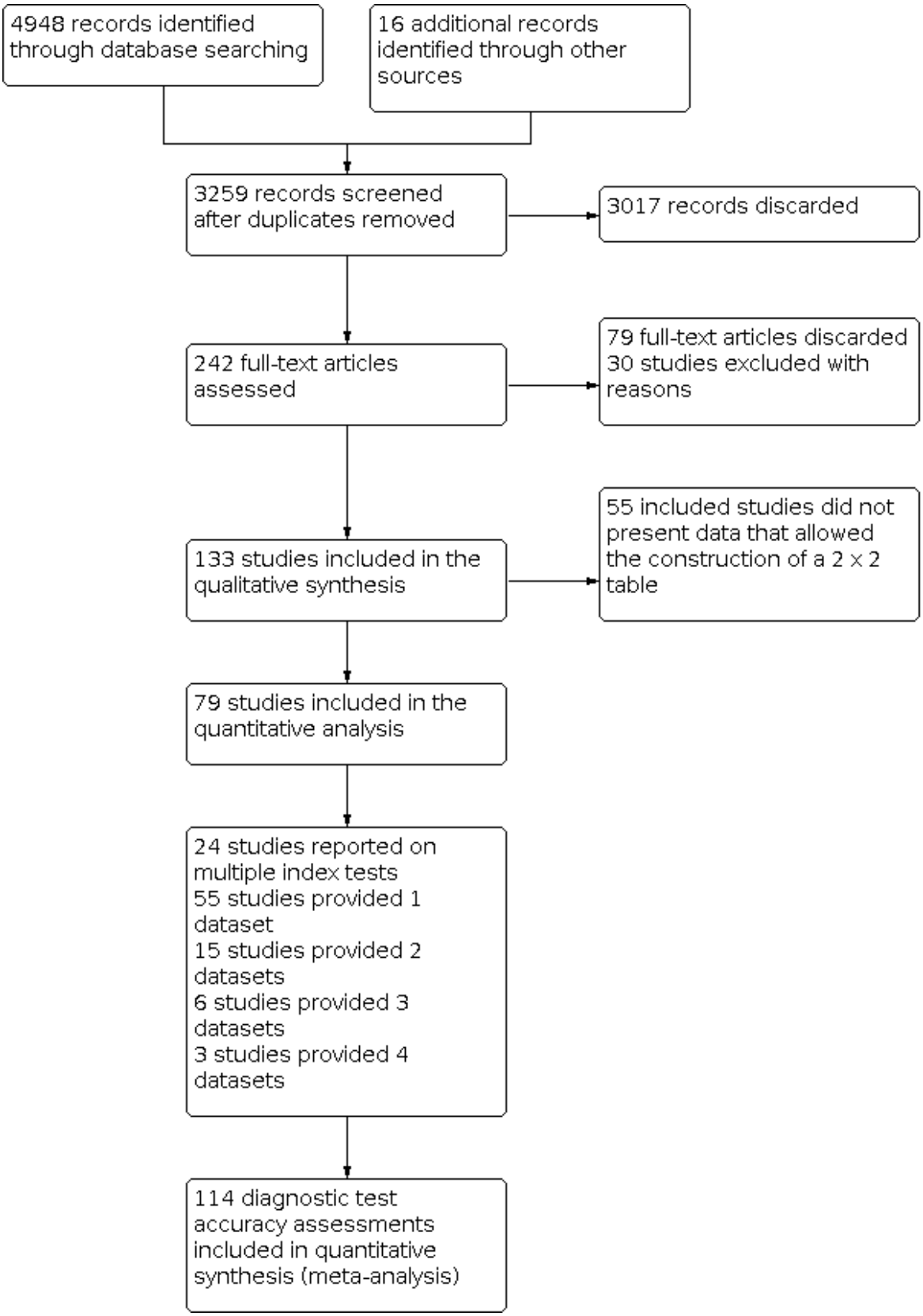
We reported our results for fluorescence index tests and the main target conditions following GRADE methodology (Schünemann 2020; Schünemann 2020a) and using the GRADEPro online tool (www.guidelinedevelopment.org). To enhance readability and understanding, we presented test accuracy results as natural frequencies to indicate numbers of false positives and false negatives. The certainty of the evidence was assessed for the overall risk of bias of the included studies, the indirectness of the evidence, the inconsistency of the results, the imprecision of the estimates, and the subjective risk of publication bias. We conducted the assessment of the certainty of the evidence irrespective of whether a numerical, a range, or a narrative description of diagnostic test accuracy was available. We categorised the certainty of the body of evidence as high, moderate, low, or very low.

RESULTS

Results of the search

The search identified a total of 3259 records after duplicates were removed. We excluded 3017 records based on the titles and abstracts, as per the eligibility criteria, the remaining 242 studies were assessed based on the full published paper. 133 of these studies were eligible for inclusion, the PRISMA diagram in Figure 2 shows the flow of studies through the review process (Moher 2009). The included studies were mainly carried out in Brazil and Europe, followed by Turkey and the Middle East, Asia, 11 in North America, and Australia. 62% (83/133) of studies performed the tests on extracted teeth, 19% in a dental school university hospital, 13% in a primary care or other clinical setting, and 6% in schools. Six of the studies (4%) reported the inclusion of fissure sealants. Studies were published between the years 1998 and 2019, 55% were published after 2010. All studies were cross-sectional and were a single gate design.

Figure 2. Review flow diagram.



Of these studies, 79 provided data in a form that allowed the construction of a 2 x 2 table and these were all included in the meta-analysis. 55 of the included 133 studies did not provide data in a format which enabled us to extract or derive true-positive, false-positive, false-negative, and true-negative results. These studies highlight the important issue of incomplete reporting of outcome data. The 79 studies that enabled extraction of data for the meta-analysis provided 114 datasets that evaluated 21,283 tooth surfaces. 21 studies included in the meta-analysis reported multiple index tests on the same tooth surfaces or participants, with eight of these investigating more than two fluorescence devices (Diniz 2011; Diniz 2012; Diniz 2019; Novaes 2012; Novaes 2016; Rodrigues 2008; Rodrigues 2011; Souza 2013). Four studies are listed twice in the meta-analysis as they investigated fluorescence devices on the primary and permanent dentition (Jablonski-Momeni 2016; Rodrigues 2009; Souza 2014) or different tooth surfaces (proximal/occlusal) (Bittar 2012). This resulted in 114 datasets included in the meta-analysis.

The authors of eight studies were contacted to request clarification on the data. Two responded providing clarity on the prevalence of disease and confirmation of the number of true-positive, false-positive, false-negative, and true-negative results; these studies were therefore included in the meta-analysis (Alomari 2015; Kockanat 2017). One author confirmed that the sample included dentinal caries and the study was therefore excluded (Menem 2017). 30 studies were excluded from this review, reasons are provided in the Characteristics of excluded studies table.

The primary objective of the systematic review and meta-analysis was to establish the diagnostic accuracy of fluorescence devices therefore all devices were initially analysed together and covariates were subsequently investigated to assess their impact. 47 of the included studies also included evaluations of other devices and were included in the other reviews in this series. An overview of these reviews compares the comparative accuracy of all the index tests under evaluation.

Of the 114 datasets in the meta-analysis, 78 were in vitro studies which assessed extracted teeth in a laboratory setting, the remaining 36 were set in dental hospitals, community settings, schools, or a primary care setting. 78 used histology as the reference standard, 25 used an enhanced visual assessment, and six relied on radiographs to provide the reference standard. Five studies used a reference standard of excavation where those teeth that were visually or radiographically determined to require restorative treatment were drilled and the severity of demineralisation confirmed. 89 assessed occlusal surfaces, 18 investigating approximal, only six reporting results on smooth surfaces, and one used the fluorescence device to assess secondary caries (sites adjacent to a prior restoration). 70 of the included studies evaluated the permanent dentition and 40 investigated the primary dentition, the remainder were either unclear or included a mixture of primary and permanent teeth. The prevalence of caries at the dentine level ranged from 0 to 0.85 and had a mean of 0.27 (standard deviation (SD) 0.17). 35 studies reported multiple assessment sites per tooth, of these 18 were included in the meta-analysis, and nine reported multiple sites on the occlusal surface (Aktan 2012; Apostolopoulou 2009; Duruturk 2011; Jablonski-Momeni 2011; Jablonski-Momeni 2012; Matos 2011; Mendes 2006; Novaes 2012a; Seremidi 2012).

The operation, positivity threshold, and interpretation of results differed according to the three categories.

- Red fluorescence: data were obtained for 84 datasets and included DIAGNOdent (46 studies), DIAGNOdent pen (34 studies), and MidWest (four studies) devices. The Canary System was not used by any included study.
    - DIAGNOdent: 46 datasets evaluated 7316 tooth sites. The device threshold that was used to determine the presence of enamel caries varied considerably between studies. The most commonly used threshold was 5, the median was 8, the minimum was 2, and the maximum value used was 20. The prevalence of dentine caries in studies included in the meta-analysis which investigated DIAGNOdent ranged from 0.03 to 0.85. 31 (65%) of the studies used histology as the reference standard, 38 (83%) assessed the occlusal surface, and 16 (37%) assessed primary teeth.
    - DIAGNOdent pen: 34 datasets evaluated 6842 tooth sites. The device threshold that determined enamel caries ranged from 3 to 28 with a median of 8, and 5 being the most commonly used threshold. The prevalence of dentine caries in studies included in the meta-analysis which investigated DIAGNOdent pen ranged from 0.01 to 0.63. 24 (71%) of the studies used histology as the reference standard, 22 (65%) assessed the occlusal surface, and 16 (50%) assessed primary teeth.
    - MidWest: four datasets evaluated 356 tooth sites. The same threshold was used across all studies, this was based on a red/green light and sound signal. The prevalence of dentine caries ranged from 0.21 to 0.63. All of the studies used histology as the reference standard and three used permanent teeth.
- Blue fluorescence: 21 datasets were included in the meta-analysis; VistaProof (18 studies), SoproLife (three studies). The Spectra caries detection device also fits into this category but no studies provided data for inclusion in the meta-analysis (Markowitz 2015).
    - VistaProof: 18 datasets evaluated 2402 sites. The device threshold used to determine enamel caries ranged from 0.90 to 1.30. The prevalence of dentine caries ranged from 0 to 0.54. 13 (72%) of the studies used histology as the reference standard, 16 (89%) assessed the occlusal surface, and four (22%) used primary teeth.
    - SoproLife: three datasets evaluated 1027 sites. The method of examination here relies on examiner interpretation of images created via the bespoke software package, therefore thresholds are not relevant to this group. The prevalence of dentine caries ranged from 0.29 to 0.68. One of the studies used histology and two used visual as the reference standard, all assessed the occlusal surface. Of the three studies, one investigated the primary dentition, one investigated the permanent dentition, and the third mixture dentition.
- Green fluorescence: often described as quantitative light-induced fluorescence (QLF) devices, were used in nine studies.
    - QLF: nine studies evaluated 3340 sites. All studies used different methods to interpret the images that were generated by the device. The prevalence of dentine caries

ranged from 0.11 to 0.63. Five datasets used histology as the reference standard (56%), six (67%) investigated occlusal surfaces, and eight (89%) used permanent teeth.

The most common reasons for exclusion from the review were studies that explicitly included participants or teeth with dentinal or frankly cavitated surfaces and were therefore ineligible. Other commonly excluded studies compared one index test with another but with no reference standard, i.e. they were comparative rather than diagnostic test accuracy studies.

A combination of visual, radiograph, and DIAGNOdent was reported in one study and this study has been reported separately (Alomari 2015).

Additional Table 3 tabulates the study characteristics for each device, the number of tooth sites, teeth, and participants evaluated, in vivo or in vitro studies, the prevalence of enamel caries (D_1), the prevalence of dentine caries (D_3), tooth surface, reference standard, and dentition.

Methodological quality of included studies

This section reports on all 133 included studies, 79 that were included in the meta-analysis, and 55 where insufficient data were provided to enable inclusion in the meta-analysis. Figure 3 summarises the results of the quality assessment of the included studies. One study could be classified as being at low risk of bias across all domains (Castilho 2016), although this study investigated third molars which were due to be extracted, and so the generalisability of the results of this study could be questioned. The results of the individual assessment of each study is provided in Figure 4.

Figure 3. Risk of bias and applicability concerns graph: review authors' judgements about each domain presented as percentages across included studies.

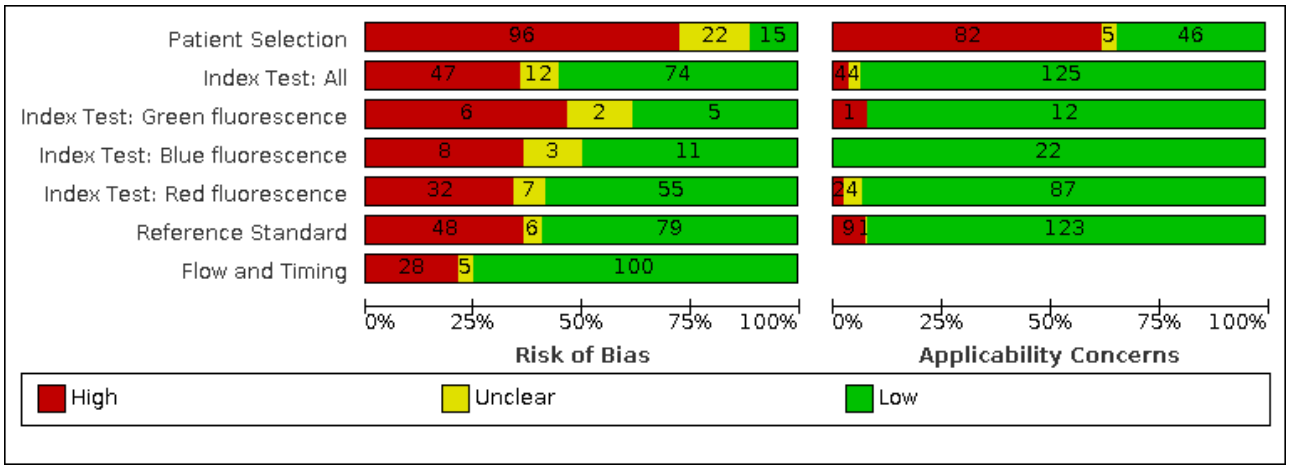


Figure 4. Risk of bias and applicability concerns summary: review authors' judgements about each domain for each included study.

Figure 4. (Continued)

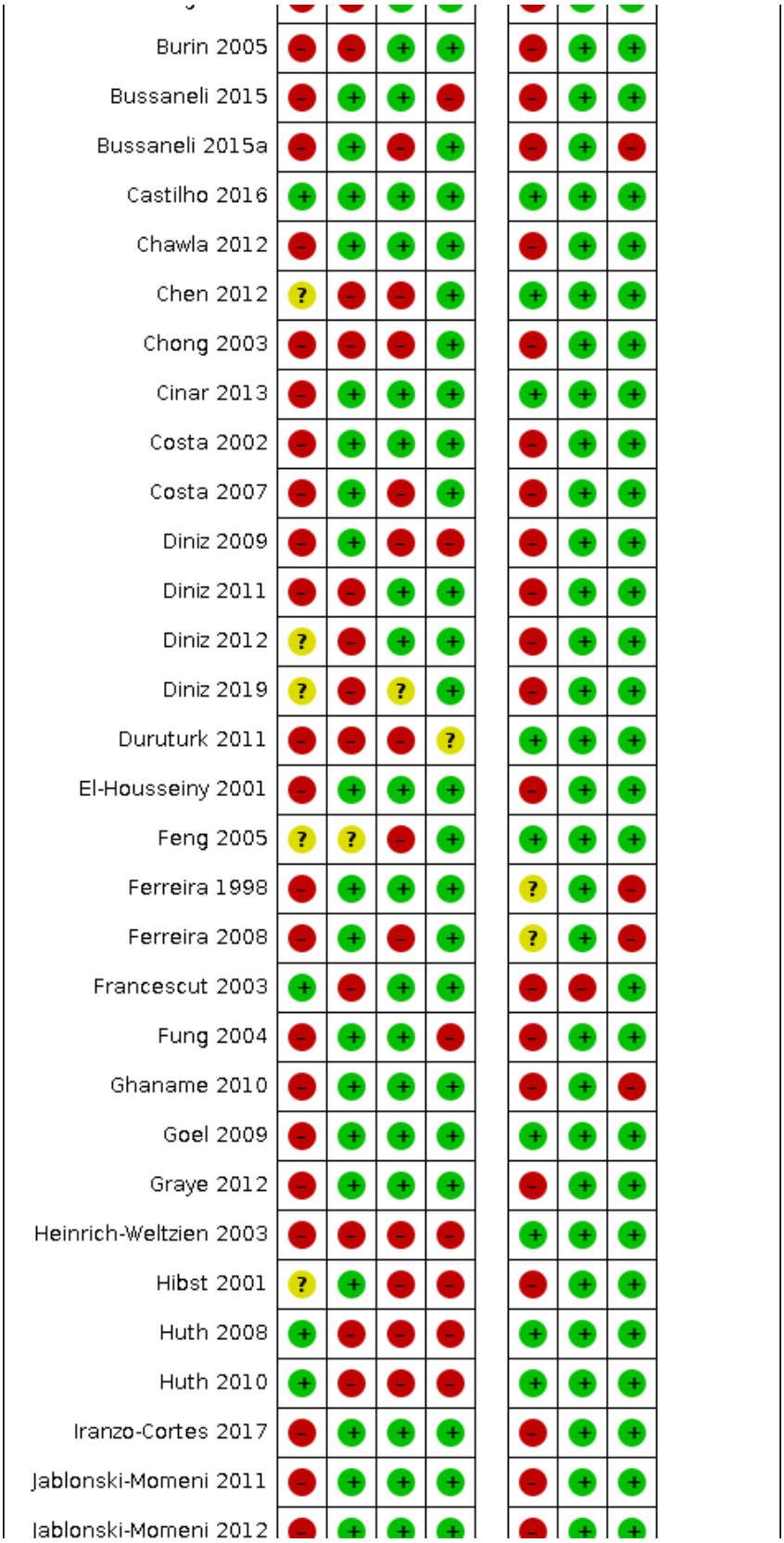


Figure 4. (Continued)

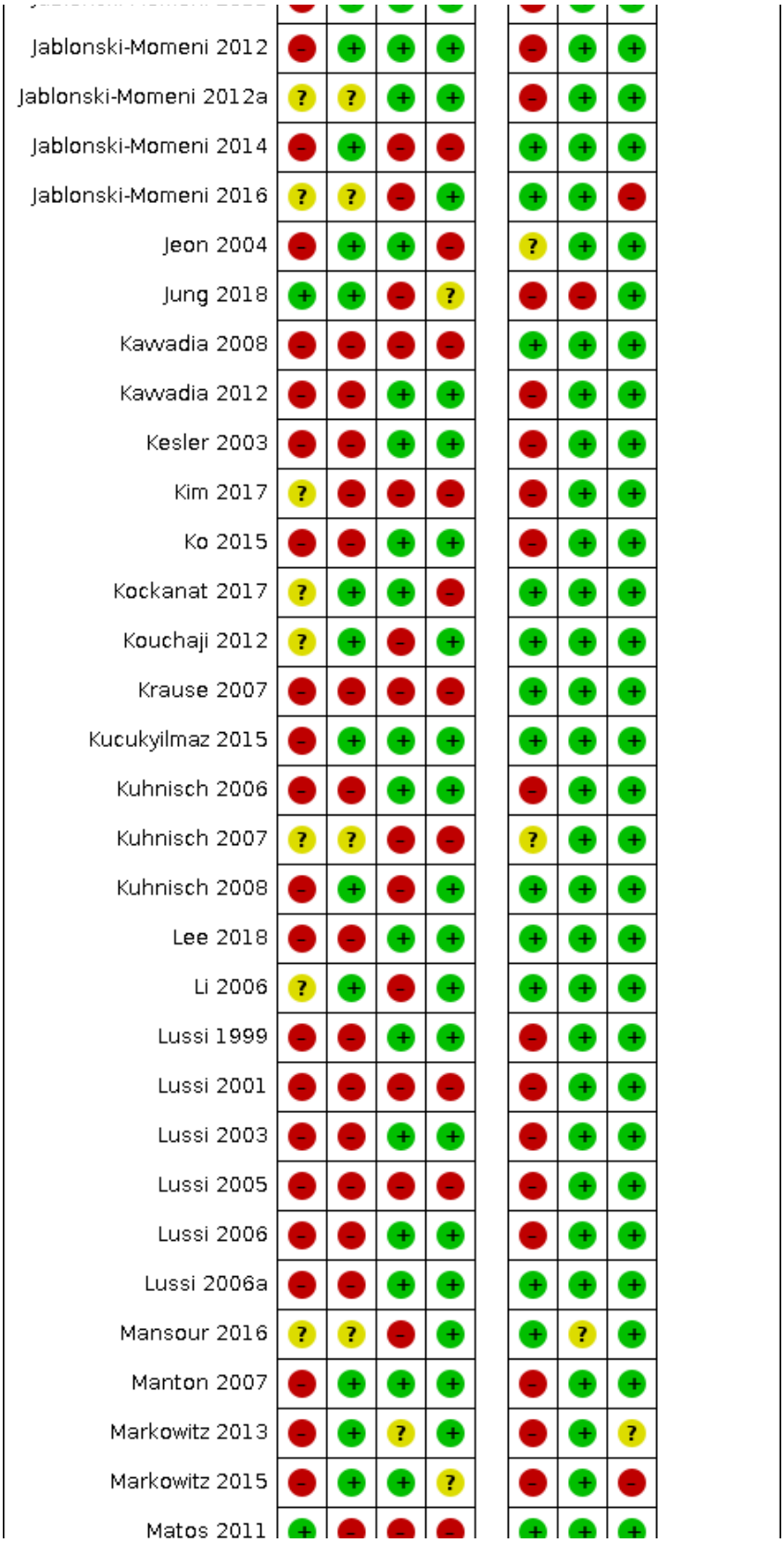


Figure 4. (Continued)

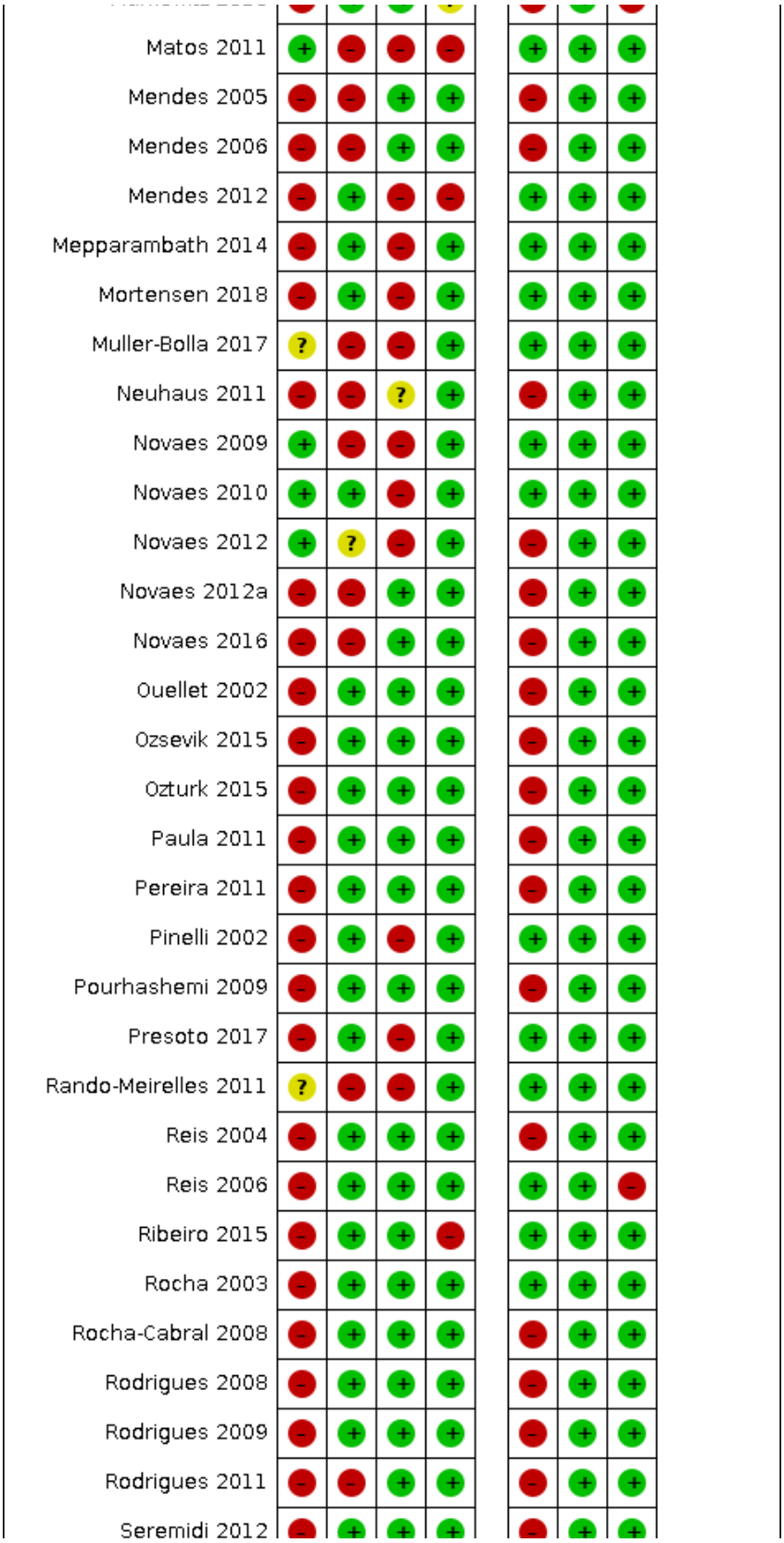
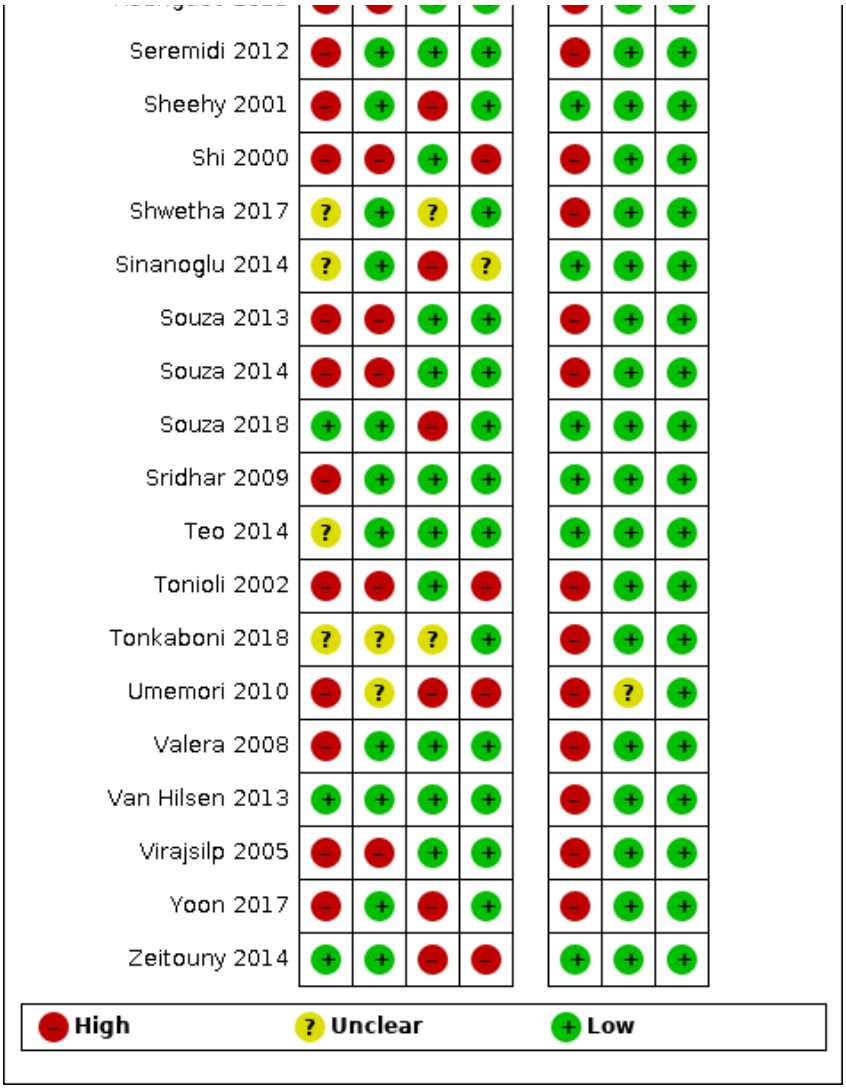


Figure 4. (Continued)



Patient selection was considered to be at low risk of bias in 15 out of 133 of studies (11%) (Almosa 2014; Anttonen 2003; Bizhang 2016; Castilho 2016; Francescut 2003; Huth 2008; Huth 2010; Jung 2018; Matos 2011; Novaes 2009; Novaes 2010; Novaes 2012; Souza 2018; Van Hilsen 2013; Zeitouny 2014), these studies clearly stated that they recruited participants or teeth consecutively or randomly. 22 of the studies (16%) failed to describe the patient selection criteria in sufficient detail and were therefore assessed as being at unclear risk of bias (Akarsu 2006; Barberia 2008; Bozdemir 2013; Chen 2012; Diniz 2012; Diniz 2019; Feng 2005; Hibst 2001; Jablonski-Momeni 2012a; Jablonski-Momeni 2016; Kim 2017; Kockanat 2017; Kouchaji 2012; Kuhnisch 2007; Li 2006; Mansour 2016; Muller-Bolla 2017; Rando-Meirelles 2011; Shwetha 2017; Sinanoglu 2014; Teo 2014; Tonkaboni 2018). The remaining 96 studies selected the participants or teeth from an available population which presented a high risk of bias to the study.

The index test was considered to be at low risk of bias in 74 out of 133 studies (55%). 47 studies (35%) were at judged as being a high risk of bias because the threshold was not pre-specified and the results of the study were used to determine the most appropriate threshold for fluorescence device.

Forty-eight studies (35%) were at high risk of bias for the reference standard. The reason for this was because the only reference standards that were accepted as correctly classifying the target condition were histology and excavation. Studies that used a reference standard of radiographs and visual examination, or a combined visual and radiograph approach as a composite reference standard, were considered to have potentially introduced bias since the target condition may not be correctly classified. 16 studies used excavation as the reference standard and there is a high level of certainty that the target condition would be observed with this method, however, the decision of whether to excavate was often based on a prior visual assessment since it would not be ethical to excavate sound or early cavitated surfaces, so the decision to allocate a high risk of bias to these studies is due to the visual or radiographic selection of teeth which were sound or had enamel caries. 79 studies used histology as the reference standard and were therefore judged at low risk of bias. There was

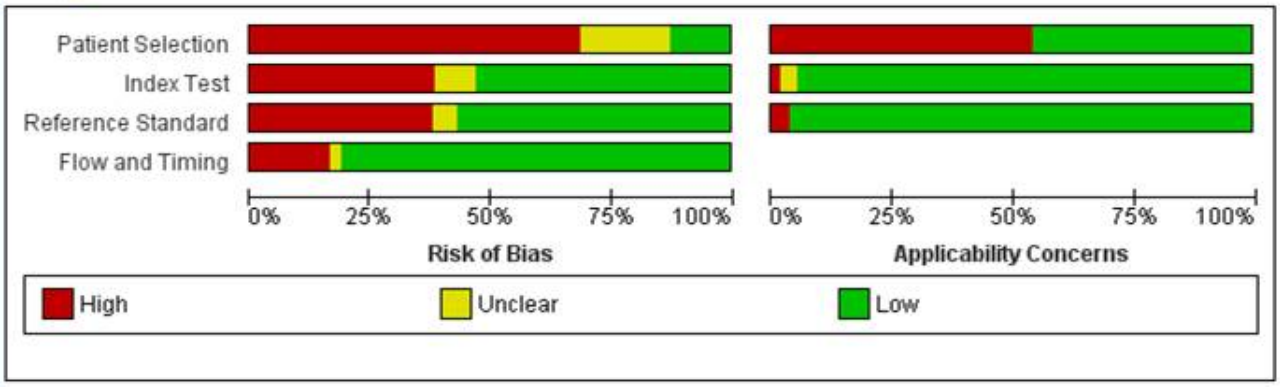
a signalling question of whether the index test results were used during the reference standard examination, this could only have occurred where the same examiner was used for index test and reference standard. Due to the teeth being extracted and sectioned for the histological examination it was decided that the results of a fluorescence device assessment would not have affected the judgement on the level of caries present, so although a negative response may be recorded for this signalling question in some cases a high risk of bias was not allocated for studies where this occurred.

Flow and timing were shown to be at high risk of bias in 28 studies (22%) (Akarsu 2006; Anttonen 2003; Bahrololoomi 2015; Bamzahim 2002; Bozdemir 2013; Bussaneli 2015; Diniz 2009; Fung 2004; Heinrich-Weltzien 2003; Hibst 2001; Huth 2008; Huth 2010; Jablonski-Momeni 2014; Jeon 2004; Kavvadia 2008; Kim 2017; Kockanat 2017; Krause 2007; Kuhnisch 2007; Lussi 2001; Lussi 2005; Matos 2011; Mendes 2012; Ribeiro 2015; Shi 2000; Tonioli 2002; Umemori 2010; Zeitouny 2014), 17 of these were because the study used a different reference standard according to the level of disease that was suspected to be present. 12 of the studies were found to be at high risk of bias for flow and timing because participants were missing from the analysis. Often this occurred because some teeth received the index test but no reference standard. If this occurred because teeth were broken during the sectioning for histological assessment and the number was explicitly reported then high risk of bias judgement was not allocated.

We assessed 82 studies (61%) as having high concern for applicability due to patient selection, these are in vitro studies where previously extracted teeth have been selected for assessment, these cannot be judged as relevant when interpreting the data for the use of devices or methods in a clinical setting. The index test was rated as a high concern for applicability in only four studies (Alomari 2015; Anttonen 2003; Francescut 2003; Jung 2018). Alomari 2015 was the only study to use a combination of visual, radiographs, and fluorescence device as the index tests, which although potentially useful to the clinician are not comparable to other included studies included in this review and was rated as not applicable. The remaining three studies used thresholds that were inappropriate, vague, or not reproducible. Four studies were unclear due to incomplete reporting of methods used to undertake the index test (Arslan 2014; Bahrololoomi 2015; Mansour 2016; Umemori 2010). The reference standard resulted in eight studies that were at high concern of applicability, this was due to a threshold being chosen that did not allow for the assessment of enamel caries.

The quality assessment and applicability of the 79 studies (Figure 5) included in the meta-analysis were compared visually to the decisions made on all 133 studies (Figure 3). We decided that the proportion of studies identified as having a high risk of bias or concern for applicability did not differ substantially between the 133 included studies and 79 studies in the meta-analysis. For example, the patient selection domain, which showed the highest proportion of high risk of bias, differed from 71% for the 133 studies to 66% for the 79 studies.

Figure 5. Studies included in the meta-analysis - Risk of bias and applicability concerns graph: review authors' judgements about each domain presented as percentages across included studies.



Findings

We evaluated the accuracy of the fluorescence devices across the 79 studies which provided 114 datasets for the meta-analysis (Figure 6 and Figure 7), the main study results are reported in Summary of findings 1. The point of assessment was the tooth surface, no studies reported at the patient level but some studies did assess multiple sites on the same surface, where this occurred it was noted in the Characteristics of included studies tables. The primary findings are reported for all available datasets with no restrictions on tooth surfaces, dentition, reference standard, or prevalence of disease. All analyses were undertaken using hierarchical summary receiver operating characteristic (HSROC) models. Observed sensitivities ranged from 0.16 to 1 and the specificities ranged from 0 to 1. The diagnostic odds ratio (DOR) was 14.12 (95% confidence interval (CI) 11.17 to 17.84). There was considerable variation in results for the different devices used, and therefore a summary sensitivity and specificity estimate has not been calculated, as a summary point on a summary receiver operating characteristic (SROC) curve estimated using mixed thresholds is clinically uninterpretable. Estimates of sensitivity and their confidence intervals were computed from the HSROC model at fixed values of specificity (median and upper quartile) to illustrate changes in sensitivity along the HSROC curve (Takwoingi 2015). At a median fixed specificity of 0.78, the estimated sensitivity was 0.70 (95% CI 0.64 to 0.75), and at an upper quartile specificity of 0.90, the sensitivity was 0.60 (95% CI 0.54 to 0.65). It should be noted that as 21 of the studies included in the meta-analysis reported the use of more than one fluorescence-based device on the same tooth surfaces, or a single fluorescence-based device on

different dentition or different tooth surfaces (proximal/occlusal), there is some non-independence of data in this analysis. No studies that directly compared tests reported the fully paired results in the form of a 2 x 4 table of the results of the index tests cross-classified amongst cases and non-cases.

Figure 6. Forest plot of all included fluorescence devices with the target condition of early/enamel caries (n = 114), ordered by sensitivity (highest to lowest).

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Achilleos 2013	36	2	0	0	1.00 [0.90, 1.00]	0.00 [0.00, 0.84]
Achilleos 2013	27	2	9	0	0.75 [0.58, 0.88]	0.00 [0.00, 0.84]
Chong 2003	48	272	0	0	1.00 [0.93, 1.00]	0.00 [0.00, 0.01]
Jablonski-Momeni 2012a	67	12	0	1	1.00 [0.95, 1.00]	0.08 [0.00, 0.36]
Seremidi 2012	81	21	2	3	0.98 [0.92, 1.00]	0.13 [0.03, 0.32]
Novaes 2016	73	26	3	7	0.96 [0.89, 0.99]	0.21 [0.09, 0.39]
Jablonski-Momeni 2011	69	19	4	6	0.95 [0.87, 0.98]	0.24 [0.09, 0.45]
Braga 2009	71	37	11	12	0.87 [0.77, 0.93]	0.24 [0.13, 0.39]
Castilho 2016	28	6	7	2	0.80 [0.63, 0.92]	0.25 [0.03, 0.65]
Jablonski-Momeni 2012	57	17	2	6	0.97 [0.88, 1.00]	0.26 [0.10, 0.48]
Diniz 2019	47	15	19	7	0.71 [0.59, 0.82]	0.32 [0.14, 0.55]
Sridhar 2009	44	4	0	2	1.00 [0.92, 1.00]	0.33 [0.04, 0.78]
Jablonski-Momeni 2014	50	157	3	96	0.94 [0.84, 0.99]	0.38 [0.32, 0.44]
Pereira 2011	53	25	2	16	0.96 [0.87, 1.00]	0.39 [0.24, 0.55]
Bittar 2012	36	10	1	8	0.97 [0.86, 1.00]	0.44 [0.22, 0.69]
Teo 2014	40	10	6	8	0.87 [0.74, 0.95]	0.44 [0.22, 0.69]
Yoon 2017	61	21	3	17	0.95 [0.87, 0.99]	0.45 [0.29, 0.62]
Novaes 2016	68	18	8	15	0.89 [0.80, 0.95]	0.45 [0.28, 0.64]
Aktan 2012	67	29	8	25	0.89 [0.80, 0.95]	0.46 [0.33, 0.60]
Diniz 2019	53	11	13	10	0.80 [0.69, 0.89]	0.48 [0.26, 0.70]
Apostolopoulou 2009	98	1	11	1	0.90 [0.83, 0.95]	0.50 [0.01, 0.99]
Diniz 2011	39	3	8	3	0.83 [0.69, 0.92]	0.50 [0.12, 0.88]
Diniz 2011	30	3	17	3	0.64 [0.49, 0.77]	0.50 [0.12, 0.88]
Goel 2009	69	1	12	1	0.85 [0.76, 0.92]	0.50 [0.01, 0.99]
Kouchaji 2012	129	11	4	12	0.97 [0.92, 0.99]	0.52 [0.31, 0.73]
Iranzo-Cortes 2017	42	7	7	8	0.86 [0.73, 0.94]	0.53 [0.27, 0.79]
Jung 2018	333	183	39	236	0.90 [0.86, 0.92]	0.56 [0.51, 0.61]
Presoto 2017	37	29	1	40	0.97 [0.86, 1.00]	0.58 [0.45, 0.70]
Heinrich-Weltzien 2003	212	10	12	14	0.95 [0.91, 0.97]	0.58 [0.37, 0.78]
Rando-Meirelles 2011	213	215	53	308	0.80 [0.75, 0.85]	0.59 [0.55, 0.63]
Souza 2018	59	45	21	70	0.74 [0.63, 0.83]	0.61 [0.51, 0.70]
Souza 2013	43	7	17	12	0.72 [0.59, 0.83]	0.63 [0.38, 0.84]
Diniz 2019	45	8	21	14	0.68 [0.56, 0.79]	0.64 [0.41, 0.83]
Bussaneli 2015a	41	6	1	11	0.98 [0.87, 1.00]	0.65 [0.38, 0.86]
Jablonski-Momeni 2012	49	8	10	15	0.83 [0.71, 0.92]	0.65 [0.43, 0.84]
Duruturk 2011	163	105	20	217	0.89 [0.84, 0.93]	0.67 [0.62, 0.72]
Ribeiro 2015	18	8	20	17	0.47 [0.31, 0.64]	0.68 [0.46, 0.85]
Lussi 2006a	89	8	4	18	0.96 [0.89, 0.99]	0.69 [0.48, 0.86]
Huth 2010	52	12	25	28	0.68 [0.56, 0.78]	0.70 [0.53, 0.83]
Rodrigues 2011	60	5	20	12	0.75 [0.64, 0.84]	0.71 [0.44, 0.90]
Akarsu 2006	112	11	15	27	0.88 [0.81, 0.93]	0.71 [0.54, 0.85]
Muller-Bolla 2017	473	60	62	148	0.88 [0.85, 0.91]	0.71 [0.64, 0.77]
Bahrololoomi 2015	88	2	14	5	0.86 [0.78, 0.92]	0.71 [0.29, 0.96]
Novaes 2012a	50	14	14	35	0.78 [0.66, 0.87]	0.71 [0.57, 0.83]
Muller-Bolla 2017	404	58	133	148	0.75 [0.71, 0.79]	0.72 [0.65, 0.78]
Pinelli 2002	78	30	31	81	0.72 [0.62, 0.80]	0.73 [0.64, 0.81]
Baseren 2003	12	5	0	14	1.00 [0.74, 1.00]	0.74 [0.49, 0.91]
Souza 2013	38	5	22	14	0.63 [0.50, 0.75]	0.74 [0.49, 0.91]
Souza 2013	44	5	16	14	0.73 [0.60, 0.84]	0.74 [0.49, 0.91]
Seremidi 2012	66	6	17	18	0.80 [0.69, 0.88]	0.75 [0.53, 0.90]
Novaes 2012a	46	12	18	37	0.72 [0.59, 0.82]	0.76 [0.61, 0.87]
Rodrigues 2011	56	4	24	13	0.70 [0.59, 0.80]	0.76 [0.50, 0.93]
Rodrigues 2011	50	4	30	13	0.63 [0.51, 0.73]	0.76 [0.50, 0.93]
Mepparambath 2014	28	31	9	101	0.76 [0.59, 0.88]	0.77 [0.68, 0.83]
Lussi 2006a	82	6	11	20	0.88 [0.80, 0.94]	0.77 [0.56, 0.91]
Aktan 2012	42	12	33	42	0.56 [0.44, 0.67]	0.78 [0.64, 0.88]
Diniz 2012	89	1	11	4	0.89 [0.81, 0.94]	0.80 [0.28, 0.99]
Diniz 2012	74	1	26	4	0.74 [0.64, 0.82]	0.80 [0.28, 0.99]
Neuhaus 2011	20	2	7	8	0.74 [0.54, 0.89]	0.80 [0.44, 0.97]
Van Hilsen 2013	16	2	16	8	0.50 [0.32, 0.68]	0.80 [0.44, 0.97]
Bamzahim 2004	26	6	8	26	0.76 [0.59, 0.89]	0.81 [0.64, 0.93]
Matos 2011	240	6	110	26	0.69 [0.63, 0.73]	0.81 [0.64, 0.93]
Novaes 2010	141	21	337	93	0.29 [0.25, 0.34]	0.82 [0.73, 0.88]
Cinar 2013	28	2	5	9	0.85 [0.68, 0.95]	0.82 [0.48, 0.98]
Diniz 2019	45	4	21	18	0.68 [0.56, 0.79]	0.82 [0.60, 0.95]
Novaes 2016	59	6	17	27	0.78 [0.67, 0.86]	0.82 [0.65, 0.93]
Bussaneli 2015	46	5	20	23	0.70 [0.57, 0.80]	0.82 [0.63, 0.94]
Attrill 2001	27	4	8	19	0.77 [0.60, 0.90]	0.83 [0.61, 0.95]
Diniz 2011	26	1	11	5	0.77 [0.62, 0.88]	0.83 [0.36, 1.00]

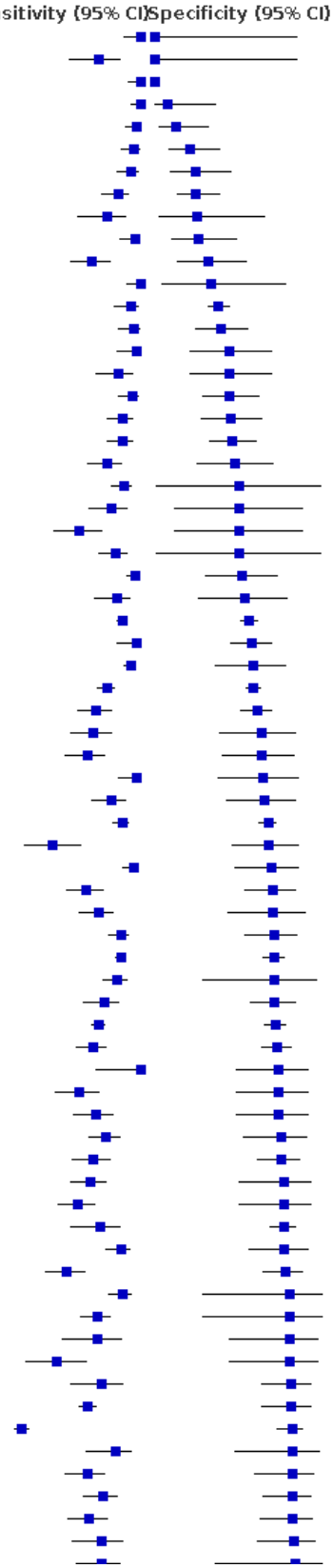


Figure 6. (Continued)

Bussaneli 2015	46	5	20	23	0.70 [0.57, 0.80]	0.82 [0.63, 0.94]
Attrill 2001	27	4	8	19	0.77 [0.60, 0.90]	0.83 [0.61, 0.95]
Diniz 2011	36	1	11	5	0.77 [0.62, 0.88]	0.83 [0.36, 1.00]
Novaes 2012a	45	8	19	41	0.70 [0.58, 0.81]	0.84 [0.70, 0.93]
Ko 2015	57	3	19	16	0.75 [0.64, 0.84]	0.84 [0.60, 0.97]
Jablonski-Momeni 2016	67	21	8	118	0.89 [0.80, 0.95]	0.85 [0.78, 0.90]
Kim 2017	72	16	98	94	0.42 [0.35, 0.50]	0.85 [0.77, 0.91]
Bussaneli 2015	44	4	22	24	0.67 [0.54, 0.78]	0.86 [0.67, 0.96]
Bittar 2012	22	3	14	19	0.61 [0.43, 0.77]	0.86 [0.65, 0.97]
Yoon 2017	35	5	29	33	0.55 [0.42, 0.67]	0.87 [0.72, 0.96]
Rodrigues 2008	76	1	35	7	0.68 [0.59, 0.77]	0.88 [0.47, 1.00]
Rodrigues 2008	99	1	12	7	0.89 [0.82, 0.94]	0.88 [0.47, 1.00]
Souza 2014	29	11	21	83	0.58 [0.43, 0.72]	0.88 [0.80, 0.94]
Zeitouny 2014	104	6	8	46	0.93 [0.86, 0.97]	0.88 [0.77, 0.96]
Pereira 2011	43	4	12	37	0.78 [0.65, 0.88]	0.90 [0.77, 0.97]
Rocha 2003	35	4	23	38	0.60 [0.47, 0.73]	0.90 [0.77, 0.97]
Matos 2011	155	3	195	29	0.44 [0.39, 0.50]	0.91 [0.75, 0.98]
Umemori 2010	25	6	11	58	0.69 [0.52, 0.84]	0.91 [0.81, 0.96]
Lee 2018	40	1	10	11	0.80 [0.66, 0.90]	0.92 [0.62, 1.00]
Rodrigues 2009	83	1	74	11	0.53 [0.45, 0.61]	0.92 [0.62, 1.00]
Rodrigues 2009	33	1	103	11	0.24 [0.17, 0.32]	0.92 [0.62, 1.00]
Lussi 2006	78	5	11	56	0.88 [0.79, 0.94]	0.92 [0.82, 0.97]
Jablonski-Momeni 2012a	49	1	22	12	0.69 [0.57, 0.79]	0.92 [0.64, 1.00]
Kockanat 2017	92	2	2	24	0.98 [0.93, 1.00]	0.92 [0.75, 0.99]
Mendes 2006	51	2	32	25	0.61 [0.50, 0.72]	0.93 [0.76, 0.99]
Ozsevik 2015	59	4	40	53	0.60 [0.49, 0.69]	0.93 [0.83, 0.98]
Jablonski-Momeni 2016	27	11	9	158	0.75 [0.58, 0.88]	0.93 [0.89, 0.97]
Costa 2002	25	1	7	16	0.78 [0.60, 0.91]	0.94 [0.71, 1.00]
Rodrigues 2011	25	1	55	16	0.31 [0.21, 0.43]	0.94 [0.71, 1.00]
Mansour 2016	36	21	15	354	0.71 [0.56, 0.83]	0.94 [0.92, 0.97]
Virajsilp 2005	67	1	22	17	0.75 [0.65, 0.84]	0.94 [0.73, 1.00]
Chen 2012	72	7	56	121	0.56 [0.47, 0.65]	0.95 [0.89, 0.98]
Feng 2005	342	75	15	1300	0.96 [0.93, 0.98]	0.95 [0.93, 0.96]
Kuhnisch 2008	489	13	110	228	0.82 [0.78, 0.85]	0.95 [0.91, 0.97]
Shi 2000	23	1	28	18	0.45 [0.31, 0.60]	0.95 [0.74, 1.00]
Mendes 2005	27	1	26	23	0.51 [0.37, 0.65]	0.96 [0.79, 1.00]
Novaes 2009	41	15	215	350	0.16 [0.12, 0.21]	0.96 [0.93, 0.98]
Souza 2014	39	2	10	51	0.80 [0.66, 0.90]	0.96 [0.87, 1.00]
Almosa 2014	317	34	223	1079	0.59 [0.54, 0.63]	0.97 [0.96, 0.98]
Kucukyilmaz 2015	107	1	57	35	0.65 [0.57, 0.72]	0.97 [0.85, 1.00]
Cinar 2013	18	0	15	11	0.55 [0.36, 0.72]	1.00 [0.72, 1.00]
Diniz 2012	85	0	15	5	0.85 [0.76, 0.91]	1.00 [0.48, 1.00]
Kockanat 2017	84	0	10	26	0.89 [0.81, 0.95]	1.00 [0.87, 1.00]
Mortensen 2018	44	0	14	2	0.76 [0.63, 0.86]	1.00 [0.16, 1.00]
Paula 2011	40	0	16	8	0.71 [0.58, 0.83]	1.00 [0.63, 1.00]
Rodrigues 2008	86	0	25	8	0.77 [0.69, 0.85]	1.00 [0.63, 1.00]
Sheehy 2001	49	0	44	77	0.53 [0.42, 0.63]	1.00 [0.95, 1.00]
Tonkaboni 2018	25	0	21	62	0.54 [0.39, 0.69]	1.00 [0.94, 1.00]

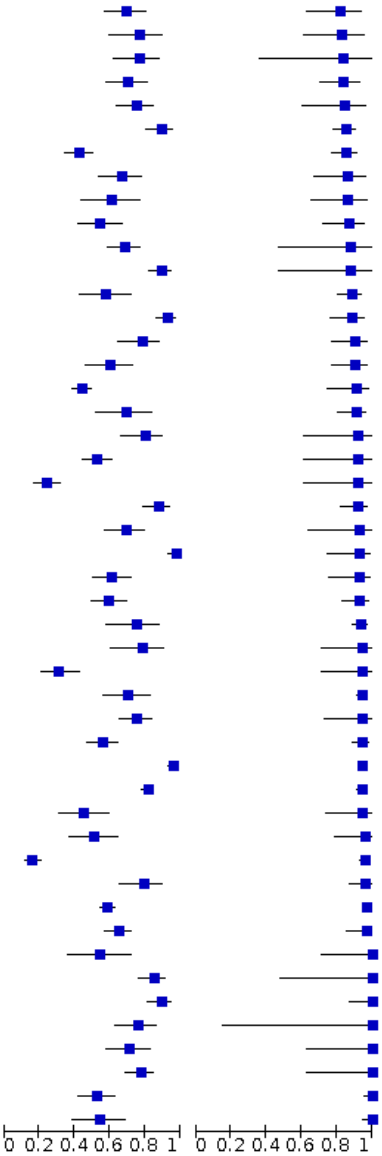
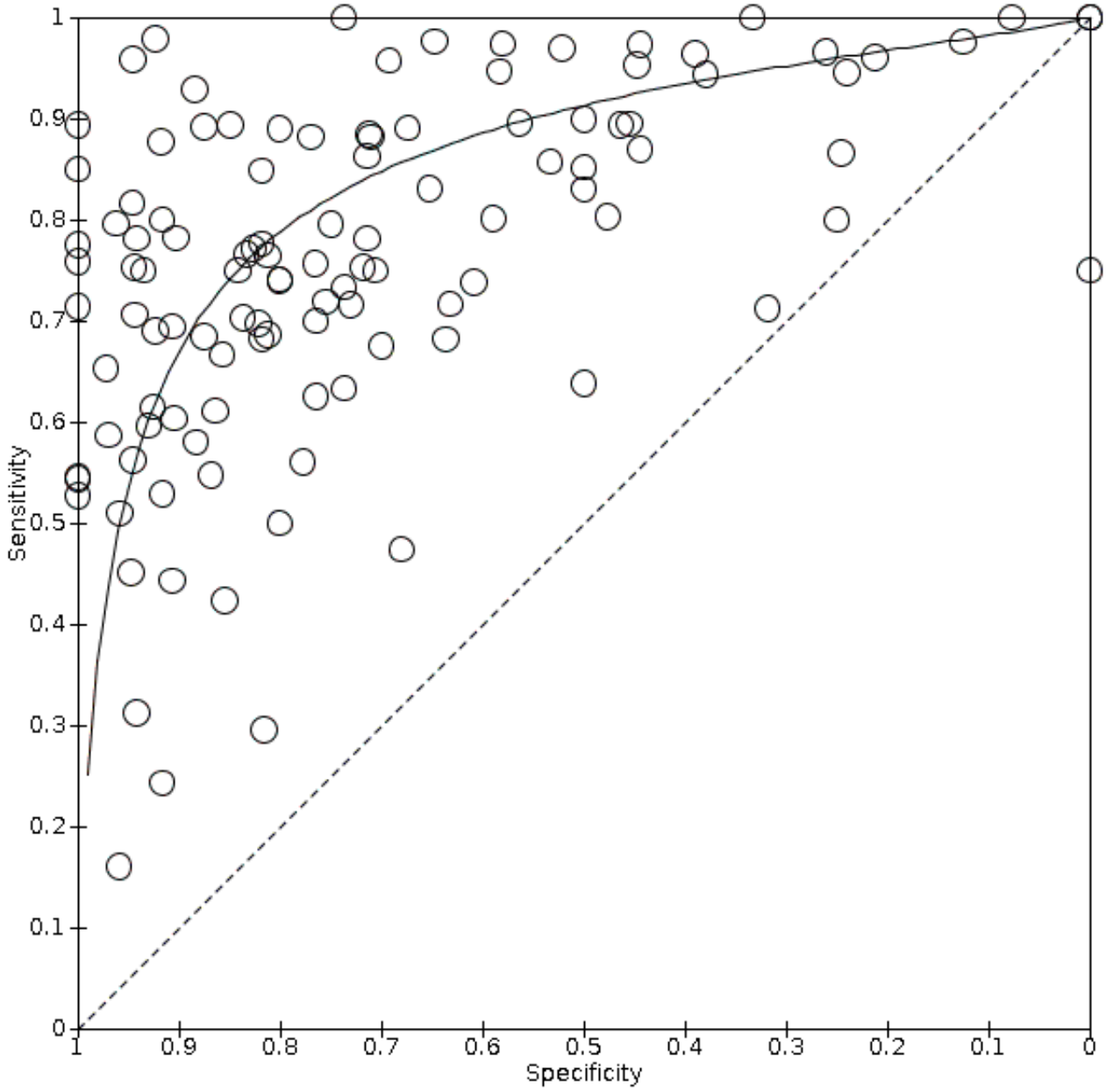


Figure 7. Summary receiver operating characteristic (SROC) plot of all fluorescence devices with the target condition of early/enamel caries (n = 114).



In accordance with the primary objective the results were categorised according to the mode of action of the devices: red, blue, or green fluorescence. We excluded one study evaluating 160 surfaces from the meta-analysis (Alomari 2015) as this study used a combined test comprising visual, radiograph, and DIAGNOdent devices. This study reported a sensitivity of 0.82 (95% CI 0.75 to 0.88) and a specificity of 0.65 (95% CI 0.38 to 0.86). The results of the meta-analysis are summarised in these subgroups in coupled forest plots (Figure 8). The HSROC model was used with covariates for device type included in the model to determine whether accuracy, threshold, or shape of the SROC curve varied with the device type. The initial, most complex model, assumed equal variances of the random effects for the different device types and included covariates to allow accuracy, threshold, and shape to vary by index test. The change in model fit was negligible when shape was removed from the model (Chi^2 = 1.89, degrees of freedom (df) = 2, P = 0.39). Finally, we explored whether all three curves took the same shape and position. The estimated HSROC curves for each of the index test categories is presented in Figure 9. We observed a visible difference between the red, blue, and green fluorescence groups which suggested that red fluorescence may be less accurate than the other two methods. However, when the covariate for accuracy was removed from the HSROC model there was only a negligible effect on the fit of the model (Chi^2 = 3.91, df = 2, P = 0.14) which indicated no statistical evidence of a difference in diagnostic accuracy according to the category of fluorescence device for caries

detection. We therefore saw no need to investigate further analyses according to these subgroups.

Figure 8. Forest plot of tests of fluorescence devices with the target condition of early/enamel caries, categorised into: red fluorescence (n = 84), blue fluorescence (n = 21), and green fluorescence (n = 9) (each group ordered by sensitivity highest to lowest).

Red fluorescence

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Baseren 2003	12	5	0	14	1.00 [0.74, 1.00]	0.74 [0.49, 0.91]
Chong 2003	48	272	0	0	1.00 [0.93, 1.00]	0.00 [0.00, 0.01]
Sridhar 2009	44	4	0	2	1.00 [0.92, 1.00]	0.33 [0.04, 0.78]
Bussaneli 2015a	41	6	1	11	0.98 [0.87, 1.00]	0.65 [0.38, 0.86]
Bittar 2012	36	10	1	8	0.97 [0.86, 1.00]	0.44 [0.22, 0.69]
Kouchaji 2012	129	11	4	12	0.97 [0.92, 0.99]	0.52 [0.31, 0.73]
Novaes 2016	73	26	3	7	0.96 [0.89, 0.99]	0.21 [0.09, 0.39]
Lussi 2006a	89	8	4	18	0.96 [0.89, 0.99]	0.69 [0.48, 0.86]
Heinrich-Weltzien 2003	212	10	12	14	0.95 [0.91, 0.97]	0.58 [0.37, 0.78]
Apostolopoulou 2009	98	1	11	1	0.90 [0.83, 0.95]	0.50 [0.01, 0.99]
Novaes 2016	68	18	8	15	0.89 [0.80, 0.95]	0.45 [0.28, 0.64]
Kockanat 2017	84	0	10	26	0.89 [0.81, 0.95]	1.00 [0.87, 1.00]
Aktan 2012	67	29	8	25	0.89 [0.80, 0.95]	0.46 [0.33, 0.60]
Rodrigues 2008	99	1	12	7	0.89 [0.82, 0.94]	0.88 [0.47, 1.00]
Duruturk 2011	163	105	20	217	0.89 [0.84, 0.93]	0.67 [0.62, 0.72]
Diniz 2012	89	1	11	4	0.89 [0.81, 0.94]	0.80 [0.28, 0.99]
Akarsu 2006	112	11	15	27	0.88 [0.81, 0.93]	0.71 [0.54, 0.85]
Lussi 2006a	82	6	11	20	0.88 [0.80, 0.94]	0.77 [0.56, 0.91]
Lussi 2006	78	5	11	56	0.88 [0.79, 0.94]	0.92 [0.82, 0.97]
Teo 2014	40	10	6	8	0.87 [0.74, 0.95]	0.44 [0.22, 0.69]
Braga 2009	71	37	11	12	0.87 [0.77, 0.93]	0.24 [0.13, 0.39]
Bahrololoomi 2015	88	2	14	5	0.86 [0.78, 0.92]	0.71 [0.29, 0.96]
Iranzo-Cortes 2017	42	7	7	8	0.86 [0.73, 0.94]	0.53 [0.27, 0.79]
Goel 2009	69	1	12	1	0.85 [0.76, 0.92]	0.50 [0.01, 0.99]
Diniz 2012	85	0	15	5	0.85 [0.76, 0.91]	1.00 [0.48, 1.00]
Cinar 2013	28	2	5	9	0.85 [0.68, 0.95]	0.82 [0.48, 0.98]
Jablonski-Momeni 2012	49	8	10	15	0.83 [0.71, 0.92]	0.65 [0.43, 0.84]
Diniz 2011	39	3	8	3	0.83 [0.69, 0.92]	0.50 [0.12, 0.88]
Kuhnisch 2008	489	13	110	228	0.82 [0.78, 0.85]	0.95 [0.91, 0.97]
Diniz 2019	53	11	13	10	0.80 [0.69, 0.89]	0.48 [0.26, 0.70]
Rando-Meirelles 2011	213	215	53	308	0.80 [0.75, 0.85]	0.59 [0.55, 0.63]
Castilho 2016	28	6	7	2	0.80 [0.63, 0.92]	0.25 [0.03, 0.65]
Souza 2014	39	2	10	51	0.80 [0.66, 0.90]	0.96 [0.87, 1.00]
Seremidi 2012	66	6	17	18	0.80 [0.69, 0.88]	0.75 [0.53, 0.90]
Pereira 2011	43	4	12	37	0.78 [0.65, 0.88]	0.90 [0.77, 0.97]
Costa 2002	25	1	7	16	0.78 [0.60, 0.91]	0.94 [0.71, 1.00]
Novaes 2012a	50	14	14	35	0.78 [0.66, 0.87]	0.71 [0.57, 0.83]
Rodrigues 2008	86	0	25	8	0.77 [0.69, 0.85]	1.00 [0.63, 1.00]
Attrill 2001	27	4	8	19	0.77 [0.60, 0.90]	0.83 [0.61, 0.95]
Bamzahim 2004	26	6	8	26	0.76 [0.59, 0.89]	0.81 [0.64, 0.93]
Mortensen 2018	44	0	14	2	0.76 [0.63, 0.86]	1.00 [0.16, 1.00]
Mepparambath 2014	28	31	9	101	0.76 [0.59, 0.88]	0.77 [0.68, 0.83]
Virajsilp 2005	67	1	22	17	0.75 [0.65, 0.84]	0.94 [0.73, 1.00]
Muller-Bolla 2017	404	58	133	148	0.75 [0.71, 0.79]	0.72 [0.65, 0.78]
Rodrigues 2011	60	5	20	12	0.75 [0.64, 0.84]	0.71 [0.44, 0.90]
Achilleos 2013	27	2	9	0	0.75 [0.58, 0.88]	0.00 [0.00, 0.84]
Neuhaus 2011	20	2	7	8	0.74 [0.54, 0.89]	0.80 [0.44, 0.97]
Souza 2018	59	45	21	70	0.74 [0.63, 0.83]	0.61 [0.51, 0.70]
Souza 2013	43	7	17	12	0.72 [0.59, 0.83]	0.63 [0.38, 0.84]
Pinelli 2002	78	30	31	81	0.72 [0.62, 0.80]	0.73 [0.64, 0.81]
Paula 2011	40	0	16	8	0.71 [0.58, 0.83]	1.00 [0.63, 1.00]
Diniz 2019	47	15	19	7	0.71 [0.59, 0.82]	0.32 [0.14, 0.55]
Mansour 2016	36	21	15	354	0.71 [0.56, 0.83]	0.94 [0.92, 0.97]
Novaes 2012a	45	8	19	41	0.70 [0.58, 0.81]	0.84 [0.70, 0.93]
Rodrigues 2011	56	4	24	13	0.70 [0.59, 0.80]	0.76 [0.50, 0.93]
Umemori 2010	25	6	11	58	0.69 [0.52, 0.84]	0.91 [0.81, 0.96]
Jablonski-Momeni 2012a	49	1	22	12	0.69 [0.57, 0.79]	0.92 [0.64, 1.00]
Matos 2011	240	6	110	26	0.69 [0.63, 0.73]	0.81 [0.64, 0.93]
Diniz 2019	45	8	21	14	0.68 [0.56, 0.79]	0.64 [0.41, 0.83]
Huth 2010	52	12	25	28	0.68 [0.56, 0.78]	0.70 [0.53, 0.83]
Bussaneli 2015	44	4	22	24	0.67 [0.54, 0.78]	0.86 [0.67, 0.96]
Kucukyilmaz 2015	107	1	57	35	0.65 [0.57, 0.72]	0.97 [0.85, 1.00]
Diniz 2011	30	3	17	3	0.64 [0.49, 0.77]	0.50 [0.12, 0.88]
Souza 2013	38	5	22	14	0.63 [0.50, 0.75]	0.74 [0.49, 0.91]
Rodrigues 2011	50	4	30	13	0.63 [0.51, 0.73]	0.76 [0.50, 0.93]
Mendes 2006	51	2	32	25	0.61 [0.50, 0.72]	0.93 [0.76, 0.99]
Bittar 2012	22	2	14	10	0.61 [0.42, 0.77]	0.86 [0.65, 0.97]

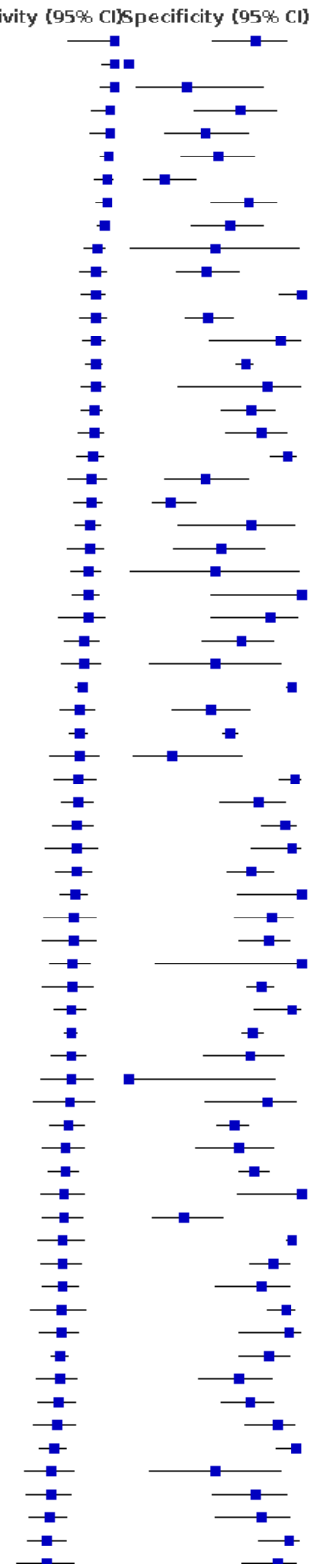
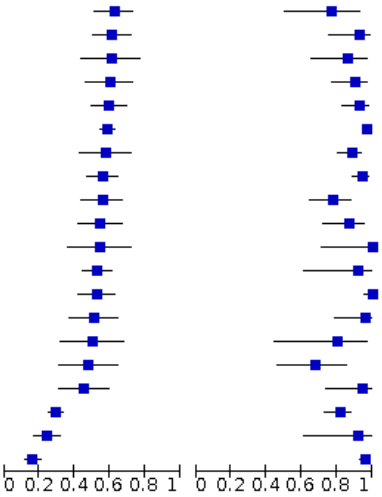


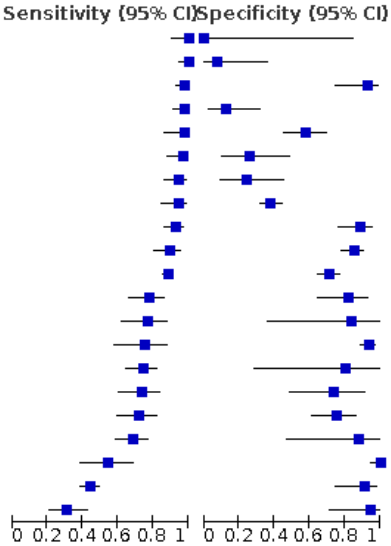
Figure 8. (Continued)

Rodrigues 2011	50	4	30	13	0.63 [0.51, 0.73]	0.76 [0.50, 0.93]
Mendes 2006	51	2	32	25	0.61 [0.50, 0.72]	0.93 [0.76, 0.99]
Bittar 2012	22	3	14	19	0.61 [0.43, 0.77]	0.86 [0.65, 0.97]
Rocha 2003	35	4	23	38	0.60 [0.47, 0.73]	0.90 [0.77, 0.97]
Ozsevik 2015	59	4	40	53	0.60 [0.49, 0.69]	0.93 [0.83, 0.98]
Almosa 2014	317	34	223	1079	0.59 [0.54, 0.63]	0.97 [0.96, 0.98]
Souza 2014	29	11	21	83	0.58 [0.43, 0.72]	0.88 [0.80, 0.94]
Chen 2012	72	7	56	121	0.56 [0.47, 0.65]	0.95 [0.89, 0.98]
Aktan 2012	42	12	33	42	0.56 [0.44, 0.67]	0.78 [0.64, 0.88]
Yoon 2017	35	5	29	33	0.55 [0.42, 0.67]	0.87 [0.72, 0.96]
Cinar 2013	18	0	15	11	0.55 [0.36, 0.72]	1.00 [0.72, 1.00]
Rodrigues 2009	83	1	74	11	0.53 [0.45, 0.61]	0.92 [0.62, 1.00]
Sheehy 2001	49	0	44	77	0.53 [0.42, 0.63]	1.00 [0.95, 1.00]
Mendes 2005	27	1	26	23	0.51 [0.37, 0.65]	0.96 [0.79, 1.00]
Van Hilsen 2013	16	2	16	8	0.50 [0.32, 0.68]	0.80 [0.44, 0.97]
Ribeiro 2015	18	8	20	17	0.47 [0.31, 0.64]	0.68 [0.46, 0.85]
Shi 2000	23	1	28	18	0.45 [0.31, 0.60]	0.95 [0.74, 1.00]
Novaes 2010	141	21	337	93	0.29 [0.25, 0.34]	0.82 [0.73, 0.88]
Rodrigues 2009	33	1	103	11	0.24 [0.17, 0.32]	0.92 [0.62, 1.00]
Novaes 2009	41	15	215	350	0.16 [0.12, 0.21]	0.96 [0.93, 0.98]



Blue fluorescence

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Achilleos 2013	36	2	0	0	1.00 [0.90, 1.00]	0.00 [0.00, 0.84]
Jablonski-Momeni 2012a	67	12	0	1	1.00 [0.95, 1.00]	0.08 [0.00, 0.36]
Kockanat 2017	92	2	2	24	0.98 [0.93, 1.00]	0.92 [0.75, 0.99]
Seremidi 2012	81	21	2	3	0.98 [0.92, 1.00]	0.13 [0.03, 0.32]
Presoto 2017	37	29	1	40	0.97 [0.86, 1.00]	0.58 [0.45, 0.70]
Jablonski-Momeni 2012	57	17	2	6	0.97 [0.88, 1.00]	0.26 [0.10, 0.48]
Jablonski-Momeni 2011	69	19	4	6	0.95 [0.87, 0.98]	0.24 [0.09, 0.45]
Jablonski-Momeni 2014	50	157	3	96	0.94 [0.84, 0.99]	0.38 [0.32, 0.44]
Zeitouny 2014	104	6	8	46	0.93 [0.86, 0.97]	0.88 [0.77, 0.96]
Jablonski-Momeni 2016	67	21	8	118	0.89 [0.80, 0.95]	0.85 [0.78, 0.90]
Muller-Bolla 2017	473	60	62	148	0.88 [0.85, 0.91]	0.71 [0.64, 0.77]
Novaes 2016	59	6	17	27	0.78 [0.67, 0.86]	0.82 [0.65, 0.93]
Diniz 2011	36	1	11	5	0.77 [0.62, 0.88]	0.83 [0.36, 1.00]
Jablonski-Momeni 2016	27	11	9	158	0.75 [0.58, 0.88]	0.93 [0.89, 0.97]
Diniz 2012	74	1	26	4	0.74 [0.64, 0.82]	0.80 [0.28, 0.99]
Souza 2013	44	5	16	14	0.73 [0.60, 0.84]	0.74 [0.49, 0.91]
Novaes 2012a	46	12	18	37	0.72 [0.59, 0.82]	0.76 [0.61, 0.87]
Rodrigues 2008	76	1	35	7	0.68 [0.59, 0.77]	0.88 [0.47, 1.00]
Tonkaboni 2018	25	0	21	62	0.54 [0.39, 0.69]	1.00 [0.94, 1.00]
Matos 2011	155	3	195	29	0.44 [0.39, 0.50]	0.91 [0.75, 0.98]
Rodrigues 2011	25	1	55	16	0.31 [0.21, 0.43]	0.94 [0.71, 1.00]



Green fluorescence

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Pereira 2011	53	25	2	16	0.96 [0.87, 1.00]	0.39 [0.24, 0.55]
Feng 2005	342	75	15	1300	0.96 [0.93, 0.98]	0.95 [0.93, 0.96]
Yoon 2017	61	21	3	17	0.95 [0.87, 0.99]	0.45 [0.29, 0.62]
Jung 2018	333	183	39	236	0.90 [0.86, 0.92]	0.56 [0.51, 0.61]
Lee 2018	40	1	10	11	0.80 [0.66, 0.90]	0.92 [0.62, 1.00]
Ko 2015	57	3	19	16	0.75 [0.64, 0.84]	0.84 [0.60, 0.97]
Bussaneli 2015	46	5	20	23	0.70 [0.57, 0.80]	0.82 [0.63, 0.94]
Diniz 2019	45	4	21	18	0.68 [0.56, 0.79]	0.82 [0.60, 0.95]
Kim 2017	72	16	98	94	0.42 [0.35, 0.50]	0.85 [0.77, 0.91]

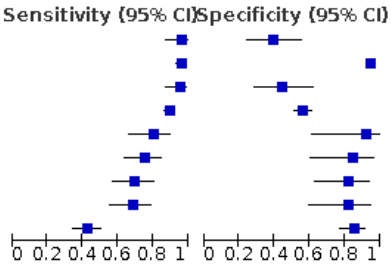
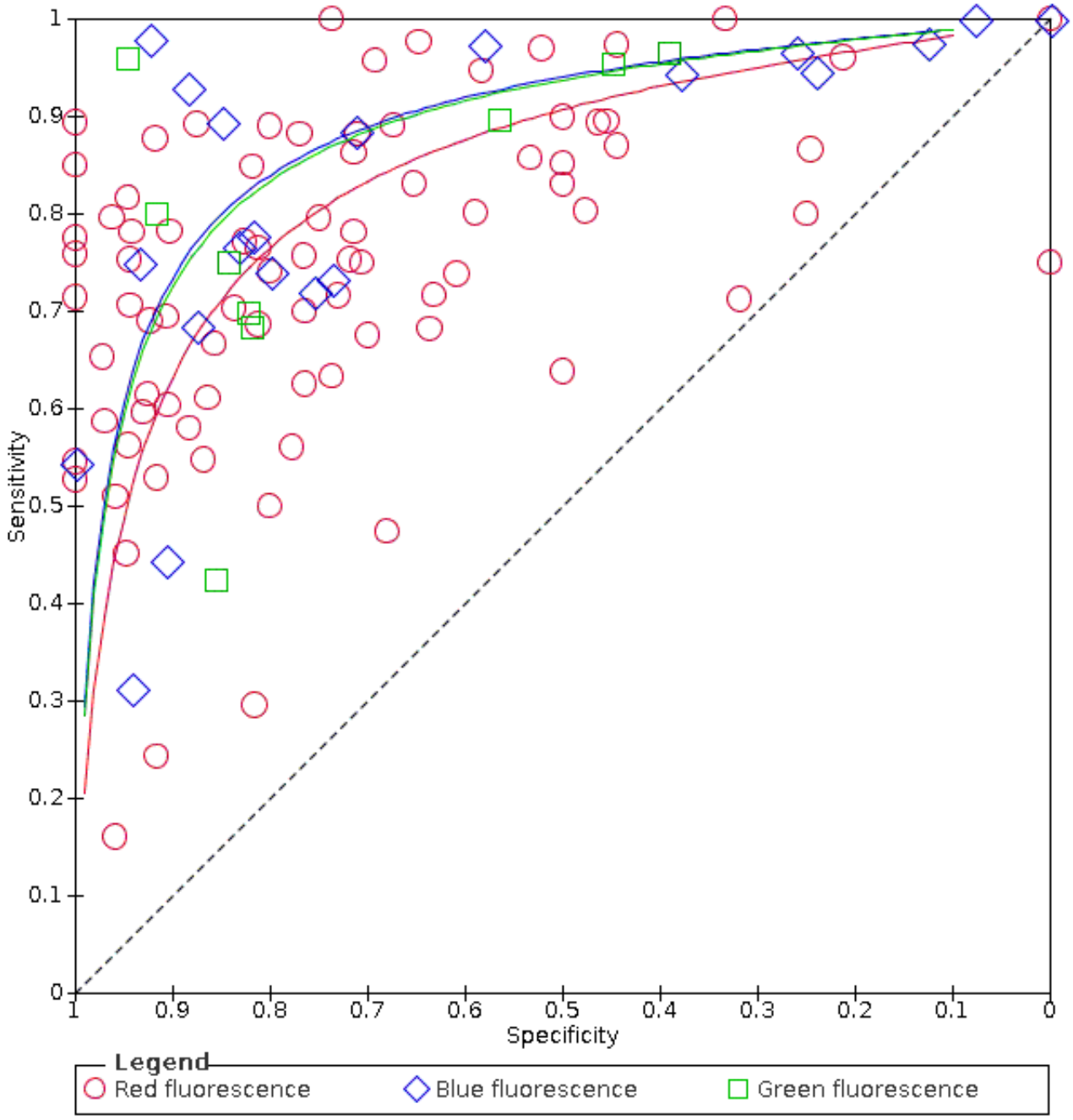


Figure 9. Summary receiver operating characteristic (SROC) plot of tests: red fluorescence (n = 84 datasets), blue fluorescence (n = 21 datasets), and green fluorescence (n = 9 datasets).



Clinically, there is interest in the performance of different devices within the three categories of fluorescence test. These have been investigated and the analyses relating to the six different devices have been included in Appendix 5.

Investigations of heterogeneity

We undertook meta-regression analysis to explore potential sources of heterogeneity. For each investigation, the initial, most complex model, assumed equal variances of the random effects for the different device types, and included covariates to allow accuracy, threshold, and shape to vary by index test. The change in model fit from the most complex model was estimated when the parameters for shape were removed from the model. Finally, the model with covariates for threshold only was estimated and compared to the model with covariates for threshold and accuracy.

Dentition

The fluorescence devices were tested on either permanent/mixed or mixed dentition. The forest plots are presented according to dentition in Figure 10 and HSROC curves were plotted for the primary and permanent groups (Figure 11). The sensitivities for permanent/mixed and primary teeth ranged from 0.31 to 1 and 0.16 to 0.98 respectively, specificities ranged from 0 to 1 and 0.09 to 1. For the purposes of analysis we combined the permanent and mixed dentition groups and compared the accuracy of the fluorescence devices on primary and permanent/mixed teeth. When covariates for dentition were included, removing shape from the model resulted in a negligible change in estimates (Chi^2 = 2.69, df = 1, P = 0.10). The accuracy of the devices on permanent/mixed dentition exceeded that of the device when used on primary teeth (Figure 11). However, when the models were tested for a difference in accuracy while leaving the shape of the curve consistent across groups there was no statistical evidence of a difference of diagnostic accuracy between the dentition (Chi^2 = 1.66, df = 1, P = 0.19). The relative diagnostic odds ratio (RDOR) for index tests on the primary dentition was 0.81 times that of tests based on permanent dentition (95% CI 0.50 to 1.31) (Additional Table 4).

Figure 10. Forest plot of datasets categorised by dentition (permanent or mixed n = 74; or primary teeth n = 40) and ordered by sensitivity.

Study	TP	FP	FN	TN	Dentition	Sensitivity (95% CI)	Specificity (95% CI)
Muller-Bolla 2017	473	60	62	148	Mixed	0.88 [0.85, 0.91]	0.71 [0.64, 0.77]
Kuhnisch 2008	489	13	110	228	Mixed	0.82 [0.78, 0.85]	0.95 [0.91, 0.97]
Rando-Meirelles 2011	213	215	53	308	Mixed	0.80 [0.75, 0.85]	0.59 [0.55, 0.63]
Muller-Bolla 2017	404	58	133	148	Mixed	0.75 [0.71, 0.79]	0.72 [0.65, 0.78]
Achilleos 2013	36	2	0	0	Permanent	1.00 [0.90, 1.00]	0.00 [0.00, 0.84]
Baseren 2003	12	5	0	14	Permanent	1.00 [0.74, 1.00]	0.74 [0.49, 0.91]
Chong 2003	48	272	0	0	Permanent	1.00 [0.93, 1.00]	0.00 [0.00, 0.01]
Jablonski-Momeni 2012a	67	12	0	1	Permanent	1.00 [0.95, 1.00]	0.08 [0.00, 0.36]
Sridhar 2009	44	4	0	2	Permanent	1.00 [0.92, 1.00]	0.33 [0.04, 0.78]
Seremidi 2012	81	21	2	3	Permanent	0.98 [0.92, 1.00]	0.13 [0.03, 0.32]
Presoto 2017	37	29	1	40	Permanent	0.97 [0.86, 1.00]	0.58 [0.45, 0.70]
Kouchaji 2012	129	11	4	12	Permanent	0.97 [0.92, 0.99]	0.52 [0.31, 0.73]
Jablonski-Momeni 2012	57	17	2	6	Permanent	0.97 [0.88, 1.00]	0.26 [0.10, 0.48]
Pereira 2011	53	25	2	16	Permanent	0.96 [0.87, 1.00]	0.39 [0.24, 0.55]
Feng 2005	342	75	15	1300	Permanent	0.96 [0.93, 0.98]	0.95 [0.93, 0.96]
Yoon 2017	61	21	3	17	Permanent	0.95 [0.87, 0.99]	0.45 [0.29, 0.62]
Heinrich-Weltzien 2003	212	10	12	14	Permanent	0.95 [0.91, 0.97]	0.58 [0.37, 0.78]
Jablonski-Momeni 2011	69	19	4	6	Permanent	0.95 [0.87, 0.98]	0.24 [0.09, 0.45]
Jablonski-Momeni 2014	50	157	3	96	Permanent	0.94 [0.84, 0.99]	0.38 [0.32, 0.44]
Zeitouny 2014	104	6	8	46	Permanent	0.93 [0.86, 0.97]	0.88 [0.77, 0.96]
Jung 2018	333	183	39	236	Permanent	0.90 [0.86, 0.92]	0.56 [0.51, 0.61]
Aktan 2012	67	29	8	25	Permanent	0.89 [0.80, 0.95]	0.46 [0.33, 0.60]
Rodrigues 2008	99	1	12	7	Permanent	0.89 [0.82, 0.94]	0.88 [0.47, 1.00]
Diniz 2012	89	1	11	4	Permanent	0.89 [0.81, 0.94]	0.80 [0.28, 0.99]
Akarsu 2006	112	11	15	27	Permanent	0.88 [0.81, 0.93]	0.71 [0.54, 0.85]
Lussi 2006	78	5	11	56	Permanent	0.88 [0.79, 0.94]	0.92 [0.82, 0.97]
Teo 2014	40	10	6	8	Permanent	0.87 [0.74, 0.95]	0.44 [0.22, 0.69]
Bahrololoomi 2015	88	2	14	5	Permanent	0.86 [0.78, 0.92]	0.71 [0.29, 0.96]
Iranzo-Cortes 2017	42	7	7	8	Permanent	0.86 [0.73, 0.94]	0.53 [0.27, 0.79]
Goel 2009	69	1	12	1	Permanent	0.85 [0.76, 0.92]	0.50 [0.01, 0.99]
Diniz 2012	85	0	15	5	Permanent	0.85 [0.76, 0.91]	1.00 [0.48, 1.00]
Jablonski-Momeni 2012	49	8	10	15	Permanent	0.83 [0.71, 0.92]	0.65 [0.43, 0.84]
Diniz 2011	39	3	8	3	Permanent	0.83 [0.69, 0.92]	0.50 [0.12, 0.88]
Castilho 2016	28	6	7	2	Permanent	0.80 [0.63, 0.92]	0.25 [0.03, 0.65]
Lee 2018	40	1	10	11	Permanent	0.80 [0.66, 0.90]	0.92 [0.62, 1.00]
Souza 2014	39	2	10	51	Permanent	0.80 [0.66, 0.90]	0.96 [0.87, 1.00]
Seremidi 2012	66	6	17	18	Permanent	0.80 [0.69, 0.88]	0.75 [0.53, 0.90]
Pereira 2011	43	4	12	37	Permanent	0.78 [0.65, 0.88]	0.90 [0.77, 0.97]
Costa 2002	25	1	7	16	Permanent	0.78 [0.60, 0.91]	0.94 [0.71, 1.00]
Rodrigues 2008	86	0	25	8	Permanent	0.77 [0.69, 0.85]	1.00 [0.63, 1.00]
Diniz 2011	36	1	11	5	Permanent	0.77 [0.62, 0.88]	0.83 [0.36, 1.00]
Bamzahim 2004	26	6	8	26	Permanent	0.76 [0.59, 0.89]	0.81 [0.64, 0.93]
Mortensen 2018	44	0	14	2	Permanent	0.76 [0.63, 0.86]	1.00 [0.16, 1.00]
Achilleos 2013	27	2	9	0	Permanent	0.75 [0.58, 0.88]	0.00 [0.00, 0.84]
Ko 2015	57	3	19	16	Permanent	0.75 [0.64, 0.84]	0.84 [0.60, 0.97]
Rodrigues 2011	60	5	20	12	Permanent	0.75 [0.64, 0.84]	0.71 [0.44, 0.90]
Diniz 2012	74	1	26	4	Permanent	0.74 [0.64, 0.82]	0.80 [0.28, 0.99]
Souza 2013	44	5	16	14	Permanent	0.73 [0.60, 0.84]	0.74 [0.49, 0.91]
Souza 2013	43	7	17	12	Permanent	0.72 [0.59, 0.83]	0.63 [0.38, 0.84]
Pinelli 2002	78	30	31	81	Permanent	0.72 [0.62, 0.80]	0.73 [0.64, 0.81]
Paula 2011	40	0	16	8	Permanent	0.71 [0.58, 0.83]	1.00 [0.63, 1.00]
Mansour 2016	36	21	15	354	Permanent	0.71 [0.56, 0.83]	0.94 [0.92, 0.97]
Rodrigues 2011	56	4	24	13	Permanent	0.70 [0.59, 0.80]	0.76 [0.50, 0.93]
Bussaneli 2015	46	5	20	23	Permanent	0.70 [0.57, 0.80]	0.82 [0.63, 0.94]
Umemori 2010	25	6	11	58	Permanent	0.69 [0.52, 0.84]	0.91 [0.81, 0.96]
Jablonski-Momeni 2012a	49	1	22	12	Permanent	0.69 [0.57, 0.79]	0.92 [0.64, 1.00]
Rodrigues 2008	76	1	35	7	Permanent	0.68 [0.59, 0.77]	0.88 [0.47, 1.00]
Huth 2010	52	12	25	28	Permanent	0.68 [0.56, 0.78]	0.70 [0.53, 0.83]
Bussaneli 2015	44	4	22	24	Permanent	0.67 [0.54, 0.78]	0.86 [0.67, 0.96]
Diniz 2011	30	3	17	3	Permanent	0.64 [0.49, 0.77]	0.50 [0.12, 0.88]
Souza 2013	38	5	22	14	Permanent	0.63 [0.50, 0.75]	0.74 [0.49, 0.91]
Rodrigues 2011	50	4	30	13	Permanent	0.63 [0.51, 0.73]	0.76 [0.50, 0.93]
Ozsevik 2015	59	4	40	53	Permanent	0.60 [0.49, 0.69]	0.93 [0.83, 0.98]
Almosa 2014	317	34	223	1079	Permanent	0.59 [0.54, 0.63]	0.97 [0.96, 0.98]
Souza 2014	29	11	21	83	Permanent	0.58 [0.43, 0.72]	0.88 [0.80, 0.94]
Aktan 2012	42	12	33	42	Permanent	0.56 [0.44, 0.67]	0.78 [0.64, 0.88]
Yoon 2017	35	5	29	33	Permanent	0.55 [0.42, 0.67]	0.87 [0.72, 0.96]
Tonkaboni 2018	25	0	21	62	Permanent	0.54 [0.39, 0.69]	1.00 [0.94, 1.00]
Rodrigues 2009	83	1	74	11	Permanent	0.53 [0.45, 0.61]	0.92 [0.62, 1.00]
Sheehy 2001	49	0	44	77	Permanent	0.53 [0.42, 0.63]	1.00 [0.95, 1.00]
Van Hilsen 2013	16	2	16	8	Permanent	0.50 [0.32, 0.68]	0.80 [0.44, 0.97]
Shi 2000	23	1	28	18	Permanent	0.45 [0.31, 0.60]	0.95 [0.74, 1.00]
Kim 2017	72	16	98	94	Permanent	0.42 [0.35, 0.50]	0.85 [0.77, 0.91]
Rodrigues 2011	25	1	55	16	Permanent	0.31 [0.21, 0.43]	0.94 [0.71, 1.00]
Rodrigues 2009	33	1	103	11	Permanent	0.24 [0.17, 0.32]	0.92 [0.62, 1.00]

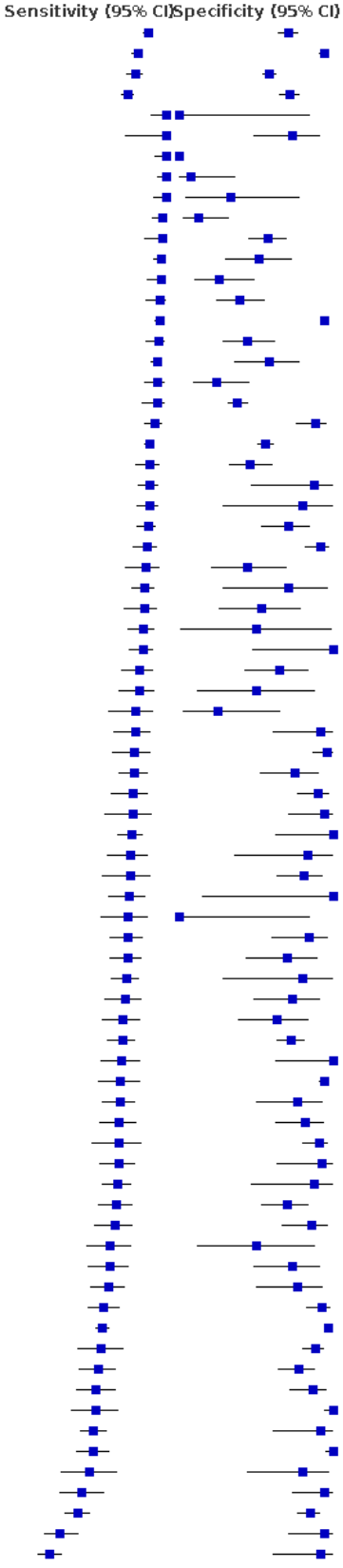


Figure 10. (Continued)

Kim 2017	72	16	98	94	Permanent	0.42 [0.35, 0.50]	0.85 [0.77, 0.91]
Rodrigues 2011	25	1	55	16	Permanent	0.31 [0.21, 0.43]	0.94 [0.71, 1.00]
Rodrigues 2009	33	1	103	11	Permanent	0.24 [0.17, 0.32]	0.92 [0.62, 1.00]
Kockanat 2017	92	2	2	24	Primary	0.98 [0.93, 1.00]	0.92 [0.75, 0.99]
Bussaneli 2015a	41	6	1	11	Primary	0.98 [0.87, 1.00]	0.65 [0.38, 0.86]
Bittar 2012	36	10	1	8	Primary	0.97 [0.86, 1.00]	0.44 [0.22, 0.69]
Novaes 2016	73	26	3	7	Primary	0.96 [0.89, 0.99]	0.21 [0.09, 0.39]
Lussi 2006a	89	8	4	18	Primary	0.96 [0.89, 0.99]	0.69 [0.48, 0.86]
Apostolopoulou 2009	98	1	11	1	Primary	0.90 [0.83, 0.95]	0.50 [0.01, 0.99]
Novaes 2016	68	18	8	15	Primary	0.89 [0.80, 0.95]	0.45 [0.28, 0.64]
Kockanat 2017	84	0	10	26	Primary	0.89 [0.81, 0.95]	1.00 [0.87, 1.00]
Jablonski-Momeni 2016	67	21	8	118	Primary	0.89 [0.80, 0.95]	0.85 [0.78, 0.90]
Duruturk 2011	163	105	20	217	Primary	0.89 [0.84, 0.93]	0.67 [0.62, 0.72]
Lussi 2006a	82	6	11	20	Primary	0.88 [0.80, 0.94]	0.77 [0.56, 0.91]
Braga 2009	71	37	11	12	Primary	0.87 [0.77, 0.93]	0.24 [0.13, 0.39]
Cinar 2013	28	2	5	9	Primary	0.85 [0.68, 0.95]	0.82 [0.48, 0.98]
Diniz 2019	53	11	13	10	Primary	0.80 [0.69, 0.89]	0.48 [0.26, 0.70]
Novaes 2012a	50	14	14	35	Primary	0.78 [0.66, 0.87]	0.71 [0.57, 0.83]
Novaes 2016	59	6	17	27	Primary	0.78 [0.67, 0.86]	0.82 [0.65, 0.93]
Attrill 2001	27	4	8	19	Primary	0.77 [0.60, 0.90]	0.83 [0.61, 0.95]
Mepparambath 2014	28	31	9	101	Primary	0.76 [0.59, 0.88]	0.77 [0.68, 0.83]
Virajsilp 2005	67	1	22	17	Primary	0.75 [0.65, 0.84]	0.94 [0.73, 1.00]
Jablonski-Momeni 2016	27	11	9	158	Primary	0.75 [0.58, 0.88]	0.93 [0.89, 0.97]
Neuhaus 2011	20	2	7	8	Primary	0.74 [0.54, 0.89]	0.80 [0.44, 0.97]
Souza 2018	59	45	21	70	Primary	0.74 [0.63, 0.83]	0.61 [0.51, 0.70]
Novaes 2012a	46	12	18	37	Primary	0.72 [0.59, 0.82]	0.76 [0.61, 0.87]
Diniz 2019	47	15	19	7	Primary	0.71 [0.59, 0.82]	0.32 [0.14, 0.55]
Novaes 2012a	45	8	19	41	Primary	0.70 [0.58, 0.81]	0.84 [0.70, 0.93]
Matos 2011	240	6	110	26	Primary	0.69 [0.63, 0.73]	0.81 [0.64, 0.93]
Diniz 2019	45	8	21	14	Primary	0.68 [0.56, 0.79]	0.64 [0.41, 0.83]
Diniz 2019	45	4	21	18	Primary	0.68 [0.56, 0.79]	0.82 [0.60, 0.95]
Kucukyilmaz 2015	107	1	57	35	Primary	0.65 [0.57, 0.72]	0.97 [0.85, 1.00]
Mendes 2006	51	2	32	25	Primary	0.61 [0.50, 0.72]	0.93 [0.76, 0.99]
Bittar 2012	22	3	14	19	Primary	0.61 [0.43, 0.77]	0.86 [0.65, 0.97]
Rocha 2003	35	4	23	38	Primary	0.60 [0.47, 0.73]	0.90 [0.77, 0.97]
Chen 2012	72	7	56	121	Primary	0.56 [0.47, 0.65]	0.95 [0.89, 0.98]
Cinar 2013	18	0	15	11	Primary	0.55 [0.36, 0.72]	1.00 [0.72, 1.00]
Mendes 2005	27	1	26	23	Primary	0.51 [0.37, 0.65]	0.96 [0.79, 1.00]
Ribeiro 2015	18	8	20	17	Primary	0.47 [0.31, 0.64]	0.68 [0.46, 0.85]
Matos 2011	155	3	195	29	Primary	0.44 [0.39, 0.50]	0.91 [0.75, 0.98]
Novaes 2010	141	21	337	93	Primary	0.29 [0.25, 0.34]	0.82 [0.73, 0.88]
Novaes 2009	41	15	215	350	Primary	0.16 [0.12, 0.21]	0.96 [0.93, 0.98]

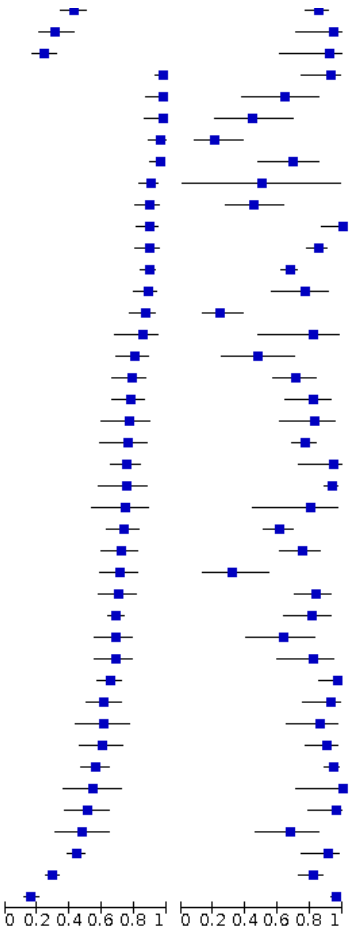
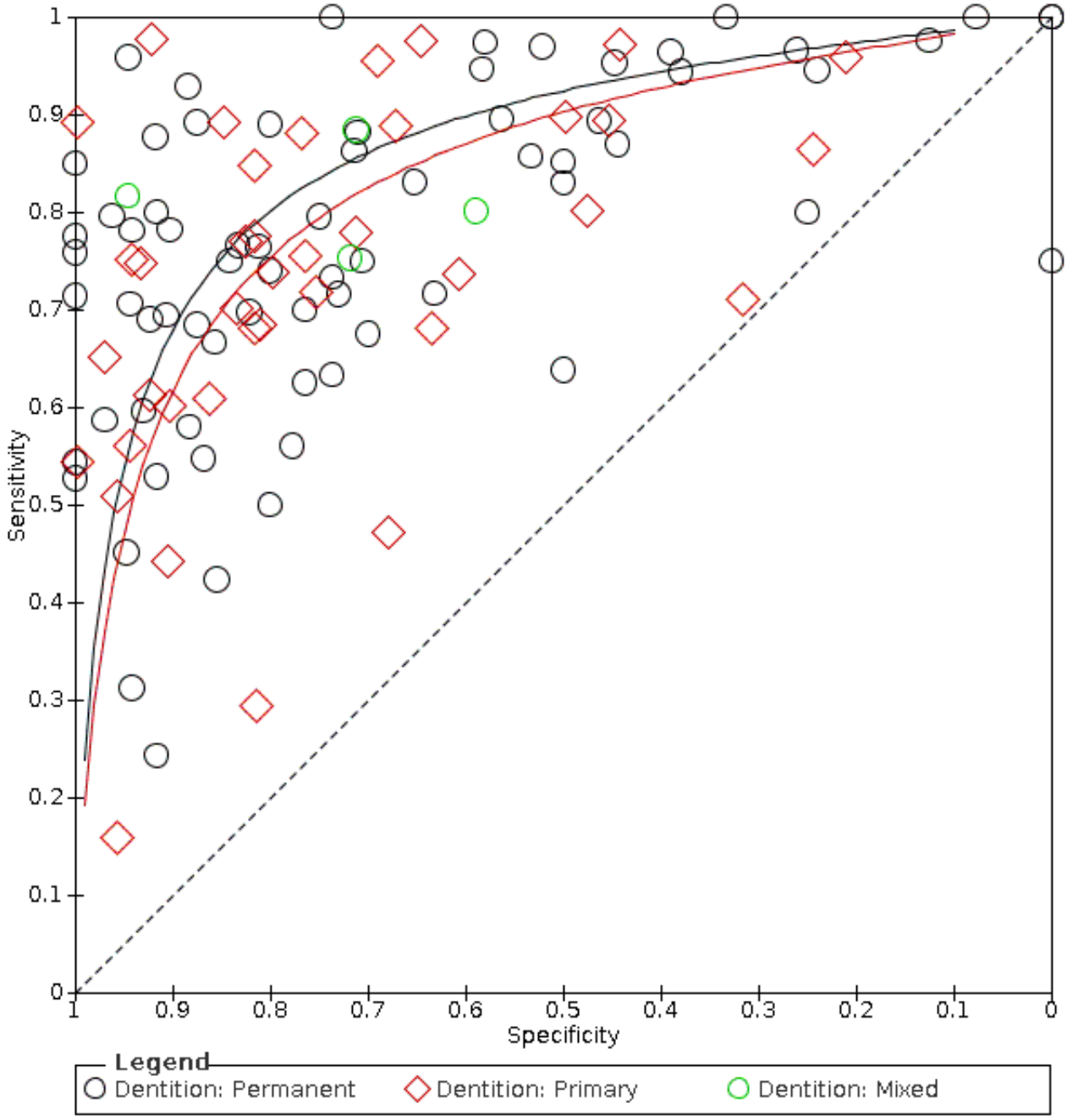


Figure 11. Summary receiver operating characteristic (SROC) plot presented according to type of dentition (permanent or mixed n = 74; or primary teeth n = 40).



Prevalence of dentine lesions

Of the 114 available datasets providing sensitivity and specificity data, the prevalence of dentine caries ranged from 0 to 0.85, with five studies not reporting the number of dentine caries in the sample (Bamzahim 2004; Feng 2005; Pinelli 2002; Presoto 2017; Yoon 2017). We created subgroups for the prevalence of dentine caries in three categories: low ≤ 14%, medium 15% to 34%, and high ≥ 35%, and for the purposes of analysis classed missing as medium prevalence; this resulted in 26 studies of low dentine prevalence, 57 medium, and 31 high. The forest plots are sorted according to the prevalence of caries into dentine in Figure 12 and the HSROC curves were plotted for the three groups (Figure 13). We observed that the estimates of sensitivity and specificity were higher for the high-prevalence datasets than the medium and low groups (Figure 13). When covariates for the prevalence of caries into dentine were included, removing shape from the model resulted in a negligible change in estimates (Chi^2 = 0.19, df = 2, P = 0.91). The accuracy of the devices on datasets with a high prevalence of dentine caries

exceeded that of low- or medium-prevalence datasets. However, when the models were tested for a difference in accuracy while leaving the shape of the curve consistent across groups there was no statistical evidence of a difference of diagnostic accuracy between the two groups (Chi^2 = 2.27, df = 2, P = 0.32). The RDOR for low prevalence was 0.76 (95% CI 0.39 to 1.48), and for medium prevalence was 1.05 (95% CI 0.59 to 1.86) (Additional Table 4) when compared with the reference category of high prevalence.

Figure 12. Forest plot of subgroups according to prevalence of dentine caries (low < 0.15, medium 0.15 to 0.34, high ≥ 0.35).

Study	TP	FP	FN	TN	Prev D3	Sensitivity (95% CI)	Specificity (95% CI)
Achilleos 2013	36	2	0	0	High	1.00 [0.90, 1.00]	0.00 [0.00, 0.84]
Jablonski-Momeni 2012a	67	12	0	1	High	1.00 [0.95, 1.00]	0.08 [0.00, 0.36]
Bussaneli 2015a	41	6	1	11	High	0.98 [0.87, 1.00]	0.65 [0.38, 0.86]
Lussi 2006a	89	8	4	18	High	0.96 [0.89, 0.99]	0.69 [0.48, 0.86]
Heinrich-Weltzien 2003	212	10	12	14	High	0.95 [0.91, 0.97]	0.58 [0.37, 0.78]
Zeitouny 2014	104	6	8	46	High	0.93 [0.86, 0.97]	0.88 [0.77, 0.96]
Rodrigues 2008	99	1	12	7	High	0.89 [0.82, 0.94]	0.88 [0.47, 1.00]
Duruturk 2011	163	105	20	217	High	0.89 [0.84, 0.93]	0.67 [0.62, 0.72]
Akarsu 2006	112	11	15	27	High	0.88 [0.81, 0.93]	0.71 [0.54, 0.85]
Lussi 2006a	82	6	11	20	High	0.88 [0.80, 0.94]	0.77 [0.56, 0.91]
Bahrololoomi 2015	88	2	14	5	High	0.86 [0.78, 0.92]	0.71 [0.29, 0.96]
Goel 2009	69	1	12	1	High	0.85 [0.76, 0.92]	0.50 [0.01, 0.99]
Diniz 2019	53	11	13	10	High	0.80 [0.69, 0.89]	0.48 [0.26, 0.70]
Rodrigues 2008	86	0	25	8	High	0.77 [0.69, 0.85]	1.00 [0.63, 1.00]
Attrill 2001	27	4	8	19	High	0.77 [0.60, 0.90]	0.83 [0.61, 0.95]
Mortensen 2018	44	0	14	2	High	0.76 [0.63, 0.86]	1.00 [0.16, 1.00]
Virajsilp 2005	67	1	22	17	High	0.75 [0.65, 0.84]	0.94 [0.73, 1.00]
Achilleos 2013	27	2	9	0	High	0.75 [0.58, 0.88]	0.00 [0.00, 0.84]
Souza 2013	44	5	16	14	High	0.73 [0.60, 0.84]	0.74 [0.49, 0.91]
Souza 2013	43	7	17	12	High	0.72 [0.59, 0.83]	0.63 [0.38, 0.84]
Diniz 2019	47	15	19	7	High	0.71 [0.59, 0.82]	0.32 [0.14, 0.55]
Jablonski-Momeni 2012a	49	1	22	12	High	0.69 [0.57, 0.79]	0.92 [0.64, 1.00]
Rodrigues 2008	76	1	35	7	High	0.68 [0.59, 0.77]	0.88 [0.47, 1.00]
Diniz 2019	45	8	21	14	High	0.68 [0.56, 0.79]	0.64 [0.41, 0.83]
Diniz 2019	45	4	21	18	High	0.68 [0.56, 0.79]	0.82 [0.60, 0.95]
Huth 2010	52	12	25	28	High	0.68 [0.56, 0.78]	0.70 [0.53, 0.83]
Souza 2013	38	5	22	14	High	0.63 [0.50, 0.75]	0.74 [0.49, 0.91]
Ozsevik 2015	59	4	40	53	High	0.60 [0.49, 0.69]	0.93 [0.83, 0.98]
Chen 2012	72	7	56	121	High	0.56 [0.47, 0.65]	0.95 [0.89, 0.98]
Tonkaboni 2018	25	0	21	62	High	0.54 [0.39, 0.69]	1.00 [0.94, 1.00]
Shi 2000	23	1	28	18	High	0.45 [0.31, 0.60]	0.95 [0.74, 1.00]
Chong 2003	48	272	0	0	Low	1.00 [0.93, 1.00]	0.00 [0.00, 0.01]
Sridhar 2009	44	4	0	2	Low	1.00 [0.92, 1.00]	0.33 [0.04, 0.78]
Jablonski-Momeni 2014	50	157	3	96	Low	0.94 [0.84, 0.99]	0.38 [0.32, 0.44]
Jung 2018	333	183	39	236	Low	0.90 [0.86, 0.92]	0.56 [0.51, 0.61]
Jablonski-Momeni 2016	67	21	8	118	Low	0.89 [0.80, 0.95]	0.85 [0.78, 0.90]
Diniz 2011	39	3	8	3	Low	0.83 [0.69, 0.92]	0.50 [0.12, 0.88]
Kuhnisch 2008	489	13	110	228	Low	0.82 [0.78, 0.85]	0.95 [0.91, 0.97]
Castilho 2016	28	6	7	2	Low	0.80 [0.63, 0.92]	0.25 [0.03, 0.65]
Lee 2018	40	1	10	11	Low	0.80 [0.66, 0.90]	0.92 [0.62, 1.00]
Diniz 2011	36	1	11	5	Low	0.77 [0.62, 0.88]	0.83 [0.36, 1.00]
Mepparambath 2014	28	31	9	101	Low	0.76 [0.59, 0.88]	0.77 [0.68, 0.83]
Jablonski-Momeni 2016	27	11	9	158	Low	0.75 [0.58, 0.88]	0.93 [0.89, 0.97]
Souza 2018	59	45	21	70	Low	0.74 [0.63, 0.83]	0.61 [0.51, 0.70]
Mansour 2016	36	21	15	354	Low	0.71 [0.56, 0.83]	0.94 [0.92, 0.97]
Umemori 2010	25	6	11	58	Low	0.69 [0.52, 0.84]	0.91 [0.81, 0.96]
Matos 2011	240	6	110	26	Low	0.69 [0.63, 0.73]	0.81 [0.64, 0.93]
Diniz 2011	30	3	17	3	Low	0.64 [0.49, 0.77]	0.50 [0.12, 0.88]
Rocha 2003	35	4	23	38	Low	0.60 [0.47, 0.73]	0.90 [0.77, 0.97]
Almosa 2014	317	34	223	1079	Low	0.59 [0.54, 0.63]	0.97 [0.96, 0.98]
Rodrigues 2009	83	1	74	11	Low	0.53 [0.45, 0.61]	0.92 [0.62, 1.00]
Mendes 2005	27	1	26	23	Low	0.51 [0.37, 0.65]	0.96 [0.79, 1.00]
Matos 2011	155	3	195	29	Low	0.44 [0.39, 0.50]	0.91 [0.75, 0.98]
Novaes 2010	141	21	337	93	Low	0.29 [0.25, 0.34]	0.82 [0.73, 0.88]
Rodrigues 2009	33	1	103	11	Low	0.24 [0.17, 0.32]	0.92 [0.62, 1.00]
Novaes 2009	41	15	215	350	Low	0.16 [0.12, 0.21]	0.96 [0.93, 0.98]
Baseren 2003	12	5	0	14	Medium	1.00 [0.74, 1.00]	0.74 [0.49, 0.91]
Kockanat 2017	92	2	2	24	Medium	0.98 [0.93, 1.00]	0.92 [0.75, 0.99]
Seremidi 2012	81	21	2	3	Medium	0.98 [0.92, 1.00]	0.13 [0.03, 0.32]
Presoto 2017	37	29	1	40	Medium	0.97 [0.86, 1.00]	0.58 [0.45, 0.70]
Bittar 2012	36	10	1	8	Medium	0.97 [0.86, 1.00]	0.44 [0.22, 0.69]
Kouchaji 2012	129	11	4	12	Medium	0.97 [0.92, 0.99]	0.52 [0.31, 0.73]
Jablonski-Momeni 2012	57	17	2	6	Medium	0.97 [0.88, 1.00]	0.26 [0.10, 0.48]
Pereira 2011	53	25	2	16	Medium	0.96 [0.87, 1.00]	0.39 [0.24, 0.55]
Novaes 2016	73	26	3	7	Medium	0.96 [0.89, 0.99]	0.21 [0.09, 0.39]
Feng 2005	342	75	15	1300	Medium	0.96 [0.93, 0.98]	0.95 [0.93, 0.96]
Yoon 2017	61	21	3	17	Medium	0.95 [0.87, 0.99]	0.45 [0.29, 0.62]
Jablonski-Momeni 2011	69	19	4	6	Medium	0.95 [0.87, 0.98]	0.24 [0.09, 0.45]
Apostolopoulou 2009	98	1	11	1	Medium	0.90 [0.83, 0.95]	0.50 [0.01, 0.99]
Novaes 2016	68	18	8	15	Medium	0.89 [0.80, 0.95]	0.45 [0.28, 0.64]
Kockanat 2017	84	0	10	26	Medium	0.89 [0.81, 0.95]	1.00 [0.87, 1.00]
Aktan 2012	67	29	8	25	Medium	0.89 [0.80, 0.95]	0.46 [0.33, 0.60]
Diniz 2012	89	1	11	4	Medium	0.89 [0.81, 0.94]	0.80 [0.28, 0.99]
Muller-Bolla 2017	473	60	62	148	Medium	0.88 [0.85, 0.91]	0.71 [0.64, 0.77]

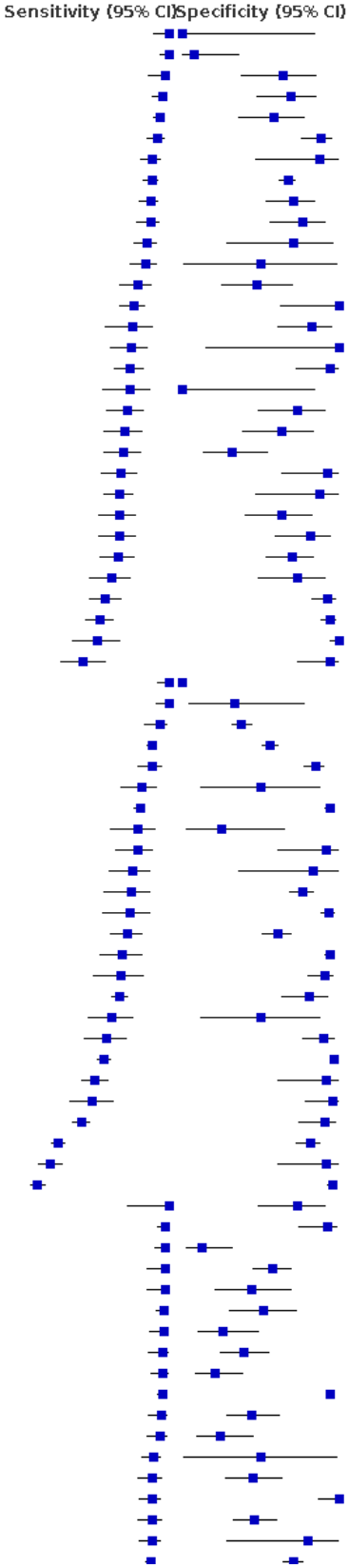


Figure 12. (Continued)

Aktan 2012	67	29	8	25	Medium	0.89 [0.80, 0.95]	0.46 [0.33, 0.60]
Diniz 2012	89	1	11	4	Medium	0.89 [0.81, 0.94]	0.80 [0.28, 0.99]
Muller-Bolla 2017	473	60	62	148	Medium	0.88 [0.85, 0.91]	0.71 [0.64, 0.77]
Lussi 2006	78	5	11	56	Medium	0.88 [0.79, 0.94]	0.92 [0.82, 0.97]
Teo 2014	40	10	6	8	Medium	0.87 [0.74, 0.95]	0.44 [0.22, 0.69]
Braga 2009	71	37	11	12	Medium	0.87 [0.77, 0.93]	0.24 [0.13, 0.39]
Iranzo-Cortes 2017	42	7	7	8	Medium	0.86 [0.73, 0.94]	0.53 [0.27, 0.79]
Diniz 2012	85	0	15	5	Medium	0.85 [0.76, 0.91]	1.00 [0.48, 1.00]
Cinar 2013	28	2	5	9	Medium	0.85 [0.68, 0.95]	0.82 [0.48, 0.98]
Jablonski-Momeni 2012	49	8	10	15	Medium	0.83 [0.71, 0.92]	0.65 [0.43, 0.84]
Rando-Meirelles 2011	213	215	53	308	Medium	0.80 [0.75, 0.85]	0.59 [0.55, 0.63]
Souza 2014	39	2	10	51	Medium	0.80 [0.66, 0.90]	0.96 [0.87, 1.00]
Seremidi 2012	66	6	17	18	Medium	0.80 [0.69, 0.88]	0.75 [0.53, 0.90]
Pereira 2011	43	4	12	37	Medium	0.78 [0.65, 0.88]	0.90 [0.77, 0.97]
Costa 2002	25	1	7	16	Medium	0.78 [0.60, 0.91]	0.94 [0.71, 1.00]
Novaes 2012a	50	14	14	35	Medium	0.78 [0.66, 0.87]	0.71 [0.57, 0.83]
Novaes 2016	59	6	17	27	Medium	0.78 [0.67, 0.86]	0.82 [0.65, 0.93]
Bamzahim 2004	26	6	8	26	Medium	0.76 [0.59, 0.89]	0.81 [0.64, 0.93]
Muller-Bolla 2017	404	58	133	148	Medium	0.75 [0.71, 0.79]	0.72 [0.65, 0.78]
Ko 2015	57	3	19	16	Medium	0.75 [0.64, 0.84]	0.84 [0.60, 0.97]
Rodrigues 2011	60	5	20	12	Medium	0.75 [0.64, 0.84]	0.71 [0.44, 0.90]
Neuhaus 2011	20	2	7	8	Medium	0.74 [0.54, 0.89]	0.80 [0.44, 0.97]
Diniz 2012	74	1	26	4	Medium	0.74 [0.64, 0.82]	0.80 [0.28, 0.99]
Novaes 2012a	46	12	18	37	Medium	0.72 [0.59, 0.82]	0.76 [0.61, 0.87]
Pinelli 2002	78	30	31	81	Medium	0.72 [0.62, 0.80]	0.73 [0.64, 0.81]
Paula 2011	40	0	16	8	Medium	0.71 [0.58, 0.83]	1.00 [0.63, 1.00]
Novaes 2012a	45	8	19	41	Medium	0.70 [0.58, 0.81]	0.84 [0.70, 0.93]
Rodrigues 2011	56	4	24	13	Medium	0.70 [0.59, 0.80]	0.76 [0.50, 0.93]
Bussaneli 2015	46	5	20	23	Medium	0.70 [0.57, 0.80]	0.82 [0.63, 0.94]
Bussaneli 2015	44	4	22	24	Medium	0.67 [0.54, 0.78]	0.86 [0.67, 0.96]
Kucukyilmaz 2015	107	1	57	35	Medium	0.65 [0.57, 0.72]	0.97 [0.85, 1.00]
Rodrigues 2011	50	4	30	13	Medium	0.63 [0.51, 0.73]	0.76 [0.50, 0.93]
Mendes 2006	51	2	32	25	Medium	0.61 [0.50, 0.72]	0.93 [0.76, 0.99]
Bittar 2012	22	3	14	19	Medium	0.61 [0.43, 0.77]	0.86 [0.65, 0.97]
Souza 2014	29	11	21	83	Medium	0.58 [0.43, 0.72]	0.88 [0.80, 0.94]
Aktan 2012	42	12	33	42	Medium	0.56 [0.44, 0.67]	0.78 [0.64, 0.88]
Yoon 2017	35	5	29	33	Medium	0.55 [0.42, 0.67]	0.87 [0.72, 0.96]
Cinar 2013	18	0	15	11	Medium	0.55 [0.36, 0.72]	1.00 [0.72, 1.00]
Sheehy 2001	49	0	44	77	Medium	0.53 [0.42, 0.63]	1.00 [0.95, 1.00]
Van Hilsen 2013	16	2	16	8	Medium	0.50 [0.32, 0.68]	0.80 [0.44, 0.97]
Ribeiro 2015	18	8	20	17	Medium	0.47 [0.31, 0.64]	0.68 [0.46, 0.85]
Kim 2017	72	16	98	94	Medium	0.42 [0.35, 0.50]	0.85 [0.77, 0.91]
Rodrigues 2011	25	1	55	16	Medium	0.31 [0.21, 0.43]	0.94 [0.71, 1.00]

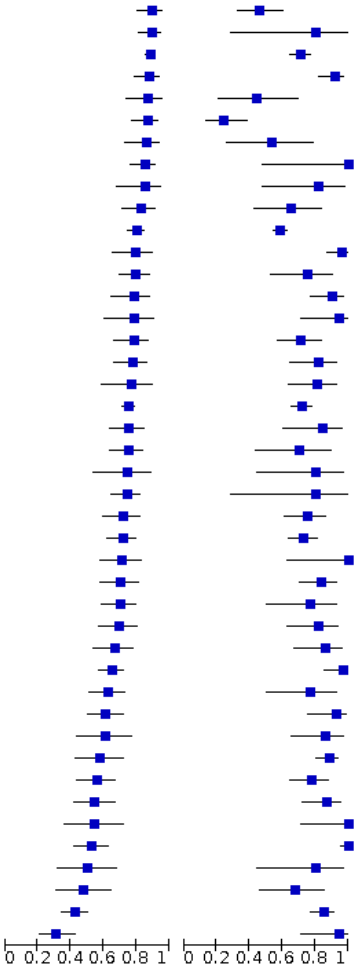
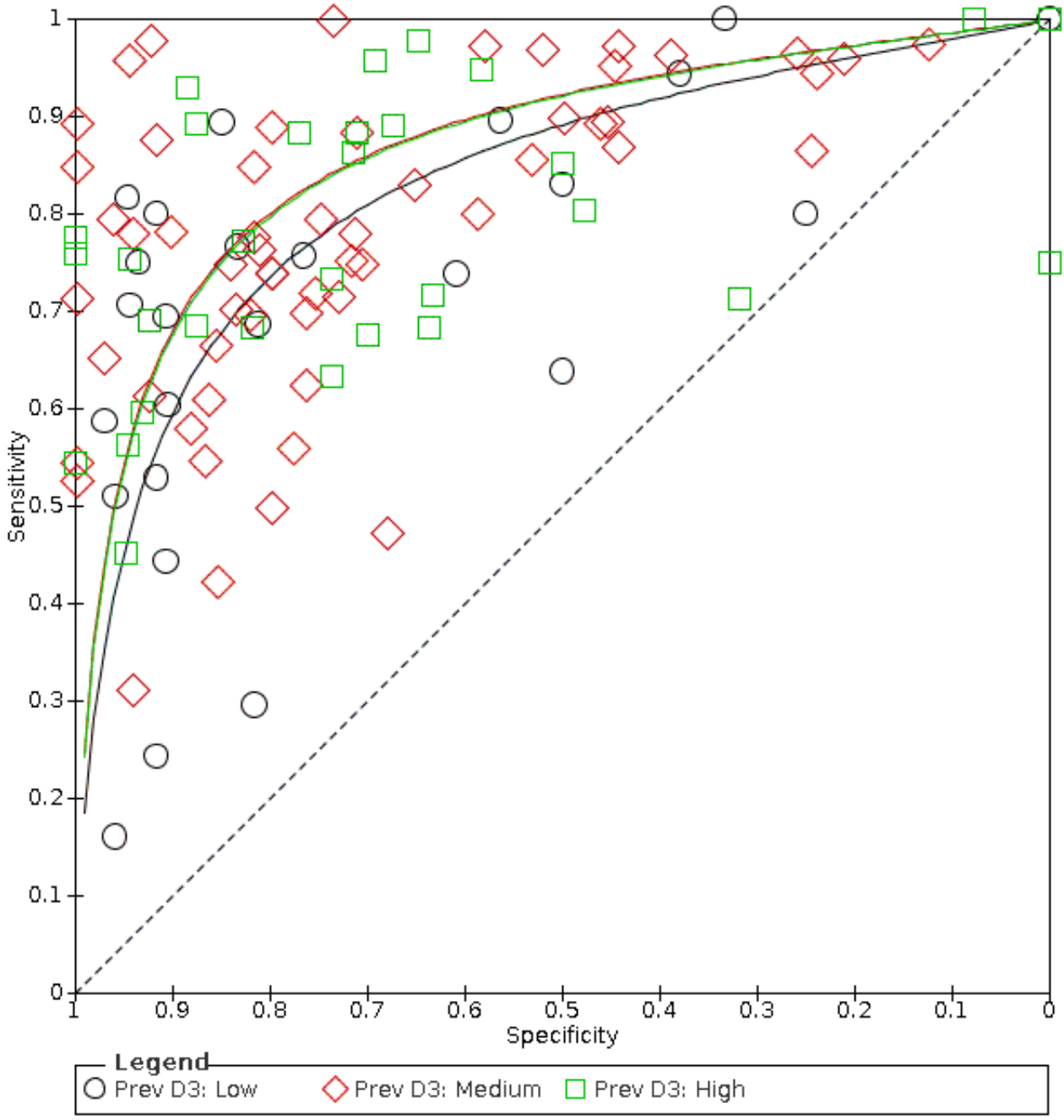


Figure 13. Summary receiver operating characteristic (SROC) plot according to prevalence of dentine caries (low < 0.15, medium 0.15 to 0.34, high ≥ 0.35).



Tooth surface

There was potential for the tooth surface to have an effect on the estimates of sensitivity and specificity. 18 datasets used the fluorescence devices on proximal surfaces, 89 datasets evaluated occlusal surfaces, and six datasets from four studies evaluated smooth surfaces (Almosa 2014; Mendes 2005; Novaes 2016; Pinelli 2002). One study focused on secondary caries and was categorised with smooth surfaces for the meta-analysis (Bamzahim 2004). Proximal, occlusal, and smooth surface results are presented as forest plots with the datasets grouped according to this covariate (Figure 14) and plotted in ROC space with HSROC curves (Figure 15). The estimates of sensitivity and specificity were higher for the occlusal and smooth surface datasets than the proximal tooth surfaces (Figure 15). When covariates for tooth surface were included, removing shape from the model resulted in a negligible change in estimates (Chi^2 = 3.29, df = 2, P = 0.19). The accuracy of the devices on datasets that evaluated occlusal datasets appeared to outperform smooth or proximal surfaces. However, when the

models were tested for a difference in accuracy while leaving the shape of the curve consistent across groups there was no statistical evidence of a difference of diagnostic accuracy between the groups (Chi^2 = 0.97, df = 2, P = 0.62). The RDOR of studies that evaluated occlusal surfaces was 1.10 (95% CI 0.59 to 2.02), and smooth/secondary caries was 1.03 (95% CI 0.36 to 2.90) compared with the reference category of proximal surfaces (Additional Table 4).

Figure 14. Forest plot of fluorescence devices according to tooth surface investigated.

Study	TP	FP	FN	TN	Surface	Sensitivity (95% CI)	Specificity (95% CI)
Bussaneli 2015a	41	6	1	11	Approximal	0.98 [0.87, 1.00]	0.65 [0.38, 0.86]
Yoon 2017	61	21	3	17	Approximal	0.95 [0.87, 0.99]	0.45 [0.29, 0.62]
Lussi 2006	78	5	11	56	Approximal	0.88 [0.79, 0.94]	0.92 [0.82, 0.97]
Souza 2014	39	2	10	51	Approximal	0.80 [0.66, 0.90]	0.96 [0.87, 1.00]
Mepparambath 2014	28	31	9	101	Approximal	0.76 [0.59, 0.88]	0.77 [0.68, 0.83]
Virajsilp 2005	67	1	22	17	Approximal	0.75 [0.65, 0.84]	0.94 [0.73, 1.00]
Ko 2015	57	3	19	16	Approximal	0.75 [0.64, 0.84]	0.84 [0.60, 0.97]
Souza 2018	59	45	21	70	Approximal	0.74 [0.63, 0.83]	0.61 [0.51, 0.70]
Ozsevik 2015	59	4	40	53	Approximal	0.60 [0.49, 0.69]	0.93 [0.83, 0.98]
Souza 2014	29	11	21	83	Approximal	0.58 [0.43, 0.72]	0.88 [0.80, 0.94]
Chen 2012	72	7	56	121	Approximal	0.56 [0.47, 0.65]	0.95 [0.89, 0.98]
Yoon 2017	35	5	29	33	Approximal	0.55 [0.42, 0.67]	0.87 [0.72, 0.96]
Tonkaboni 2018	25	0	21	62	Approximal	0.54 [0.39, 0.69]	1.00 [0.94, 1.00]
Ribeiro 2015	18	8	20	17	Approximal	0.47 [0.31, 0.64]	0.68 [0.46, 0.85]
Kim 2017	72	16	98	94	Approximal	0.42 [0.35, 0.50]	0.85 [0.77, 0.91]
Novaes 2010	141	21	337	93	Approximal	0.29 [0.25, 0.34]	0.82 [0.73, 0.88]
Novaes 2009	41	15	215	350	Approximal	0.16 [0.12, 0.21]	0.96 [0.93, 0.98]
Achilleos 2013	36	2	0	0	Occlusal	1.00 [0.90, 1.00]	0.00 [0.00, 0.84]
Baseren 2003	12	5	0	14	Occlusal	1.00 [0.74, 1.00]	0.74 [0.49, 0.91]
Chong 2003	48	272	0	0	Occlusal	1.00 [0.93, 1.00]	0.00 [0.00, 0.01]
Jablonski-Momeni 2012a	67	12	0	1	Occlusal	1.00 [0.95, 1.00]	0.08 [0.00, 0.36]
Sridhar 2009	44	4	0	2	Occlusal	1.00 [0.92, 1.00]	0.33 [0.04, 0.78]
Kockanat 2017	92	2	2	24	Occlusal	0.98 [0.93, 1.00]	0.92 [0.75, 0.99]
Seremidi 2012	81	21	2	3	Occlusal	0.98 [0.92, 1.00]	0.13 [0.03, 0.32]
Presoto 2017	37	29	1	40	Occlusal	0.97 [0.86, 1.00]	0.58 [0.45, 0.70]
Bittar 2012	36	10	1	8	Occlusal	0.97 [0.86, 1.00]	0.44 [0.22, 0.69]
Kouchaji 2012	129	11	4	12	Occlusal	0.97 [0.92, 0.99]	0.52 [0.31, 0.73]
Jablonski-Momeni 2012	57	17	2	6	Occlusal	0.97 [0.88, 1.00]	0.26 [0.10, 0.48]
Pereira 2011	53	25	2	16	Occlusal	0.96 [0.87, 1.00]	0.39 [0.24, 0.55]
Feng 2005	342	75	15	1300	Occlusal	0.96 [0.93, 0.98]	0.95 [0.93, 0.96]
Lussi 2006a	89	8	4	18	Occlusal	0.96 [0.89, 0.99]	0.69 [0.48, 0.86]
Heinrich-Weltzien 2003	212	10	12	14	Occlusal	0.95 [0.91, 0.97]	0.58 [0.37, 0.78]
Jablonski-Momeni 2011	69	19	4	6	Occlusal	0.95 [0.87, 0.98]	0.24 [0.09, 0.45]
Jablonski-Momeni 2014	50	157	3	96	Occlusal	0.94 [0.84, 0.99]	0.38 [0.32, 0.44]
Zeitouny 2014	104	6	8	46	Occlusal	0.93 [0.86, 0.97]	0.88 [0.77, 0.96]
Apostolopoulou 2009	98	1	11	1	Occlusal	0.90 [0.83, 0.95]	0.50 [0.01, 0.99]
Jung 2018	333	183	39	236	Occlusal	0.90 [0.86, 0.92]	0.56 [0.51, 0.61]
Kockanat 2017	84	0	10	26	Occlusal	0.89 [0.81, 0.95]	1.00 [0.87, 1.00]
Aktan 2012	67	29	8	25	Occlusal	0.89 [0.80, 0.95]	0.46 [0.33, 0.60]
Jablonski-Momeni 2016	67	21	8	118	Occlusal	0.89 [0.80, 0.95]	0.85 [0.78, 0.90]
Rodrigues 2008	99	1	12	7	Occlusal	0.89 [0.82, 0.94]	0.88 [0.47, 1.00]
Duruturk 2011	163	105	20	217	Occlusal	0.89 [0.84, 0.93]	0.67 [0.62, 0.72]
Diniz 2012	89	1	11	4	Occlusal	0.89 [0.81, 0.94]	0.80 [0.28, 0.99]
Muller-Bolla 2017	473	60	62	148	Occlusal	0.88 [0.85, 0.91]	0.71 [0.64, 0.77]
Akarsu 2006	112	11	15	27	Occlusal	0.88 [0.81, 0.93]	0.71 [0.54, 0.85]
Lussi 2006a	82	6	11	20	Occlusal	0.88 [0.80, 0.94]	0.77 [0.56, 0.91]
Teo 2014	40	10	6	8	Occlusal	0.87 [0.74, 0.95]	0.44 [0.22, 0.69]
Braga 2009	71	37	11	12	Occlusal	0.87 [0.77, 0.93]	0.24 [0.13, 0.39]
Bahrololoomi 2015	88	2	14	5	Occlusal	0.86 [0.78, 0.92]	0.71 [0.29, 0.96]
Iranzo-Cortes 2017	42	7	7	8	Occlusal	0.86 [0.73, 0.94]	0.53 [0.27, 0.79]
Goel 2009	69	1	12	1	Occlusal	0.85 [0.76, 0.92]	0.50 [0.01, 0.99]
Diniz 2012	85	0	15	5	Occlusal	0.85 [0.76, 0.91]	1.00 [0.48, 1.00]
Cinar 2013	28	2	5	9	Occlusal	0.85 [0.68, 0.95]	0.82 [0.48, 0.98]
Jablonski-Momeni 2012	49	8	10	15	Occlusal	0.83 [0.71, 0.92]	0.65 [0.43, 0.84]
Diniz 2011	39	3	8	3	Occlusal	0.83 [0.69, 0.92]	0.50 [0.12, 0.88]
Kuhnisch 2008	489	13	110	228	Occlusal	0.82 [0.78, 0.85]	0.95 [0.91, 0.97]
Diniz 2019	53	11	13	10	Occlusal	0.80 [0.69, 0.89]	0.48 [0.26, 0.70]
Rando-Meirelles 2011	213	215	53	308	Occlusal	0.80 [0.75, 0.85]	0.59 [0.55, 0.63]
Castilho 2016	28	6	7	2	Occlusal	0.80 [0.63, 0.92]	0.25 [0.03, 0.65]
Lee 2018	40	1	10	11	Occlusal	0.80 [0.66, 0.90]	0.92 [0.62, 1.00]
Seremidi 2012	66	6	17	18	Occlusal	0.80 [0.69, 0.88]	0.75 [0.53, 0.90]
Pereira 2011	43	4	12	37	Occlusal	0.78 [0.65, 0.88]	0.90 [0.77, 0.97]
Costa 2002	25	1	7	16	Occlusal	0.78 [0.60, 0.91]	0.94 [0.71, 1.00]
Novaes 2012a	50	14	14	35	Occlusal	0.78 [0.66, 0.87]	0.71 [0.57, 0.83]
Rodrigues 2008	86	0	25	8	Occlusal	0.77 [0.69, 0.85]	1.00 [0.63, 1.00]
Attrill 2001	27	4	8	19	Occlusal	0.77 [0.60, 0.90]	0.83 [0.61, 0.95]
Diniz 2011	36	1	11	5	Occlusal	0.77 [0.62, 0.88]	0.83 [0.36, 1.00]
Mortensen 2018	44	0	14	2	Occlusal	0.76 [0.63, 0.86]	1.00 [0.16, 1.00]
Muller-Bolla 2017	404	58	133	148	Occlusal	0.75 [0.71, 0.79]	0.72 [0.65, 0.78]
Achilleos 2013	27	2	9	0	Occlusal	0.75 [0.58, 0.88]	0.00 [0.00, 0.84]
Jablonski-Momeni 2016	27	11	9	158	Occlusal	0.75 [0.58, 0.88]	0.93 [0.89, 0.97]
Rodrigues 2011	60	5	20	12	Occlusal	0.75 [0.64, 0.84]	0.71 [0.44, 0.90]
Neuhaus 2011	20	2	7	8	Occlusal	0.74 [0.54, 0.89]	0.80 [0.44, 0.97]
Diniz 2012	74	1	26	4	Occlusal	0.74 [0.64, 0.82]	0.80 [0.28, 0.99]
Souza 2013	44	5	16	14	Occlusal	0.73 [0.60, 0.84]	0.74 [0.49, 0.91]

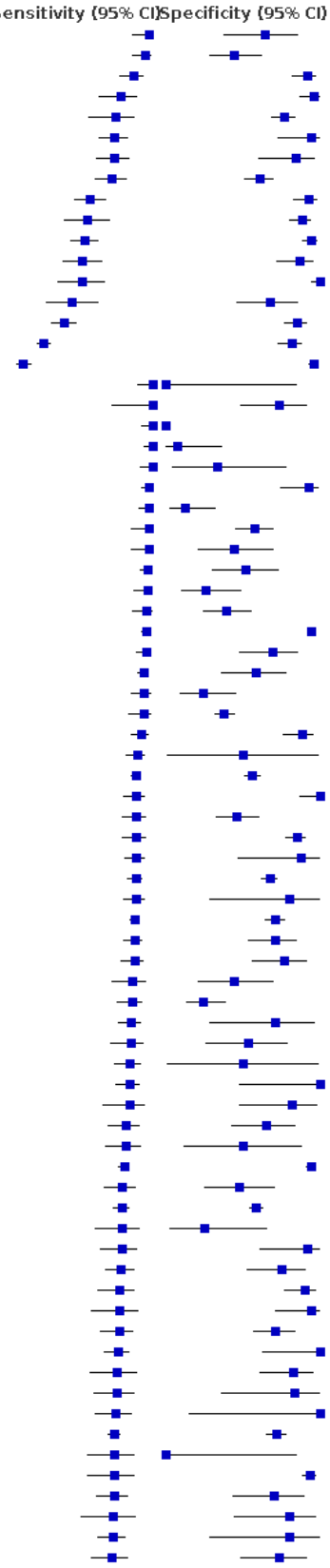


Figure 14. (Continued)

Neuhaus 2011	20	2	7	8	Occlusal	0.74 [0.54, 0.89]	0.80 [0.44, 0.97]
Diniz 2012	74	1	26	4	Occlusal	0.74 [0.64, 0.82]	0.80 [0.28, 0.99]
Souza 2013	44	5	16	14	Occlusal	0.73 [0.60, 0.84]	0.74 [0.49, 0.91]
Novaes 2012a	46	12	18	37	Occlusal	0.72 [0.59, 0.82]	0.76 [0.61, 0.87]
Souza 2013	43	7	17	12	Occlusal	0.72 [0.59, 0.83]	0.63 [0.38, 0.84]
Paula 2011	40	0	16	8	Occlusal	0.71 [0.58, 0.83]	1.00 [0.63, 1.00]
Diniz 2019	47	15	19	7	Occlusal	0.71 [0.59, 0.82]	0.32 [0.14, 0.55]
Mansour 2016	36	21	15	354	Occlusal	0.71 [0.56, 0.83]	0.94 [0.92, 0.97]
Novaes 2012a	45	8	19	41	Occlusal	0.70 [0.58, 0.81]	0.84 [0.70, 0.93]
Rodrigues 2011	56	4	24	13	Occlusal	0.70 [0.59, 0.80]	0.76 [0.50, 0.93]
Bussaneli 2015	46	5	20	23	Occlusal	0.70 [0.57, 0.80]	0.82 [0.63, 0.94]
Umemori 2010	25	6	11	58	Occlusal	0.69 [0.52, 0.84]	0.91 [0.81, 0.96]
Jablonski-Momeni 2012a	49	1	22	12	Occlusal	0.69 [0.57, 0.79]	0.92 [0.64, 1.00]
Matos 2011	240	6	110	26	Occlusal	0.69 [0.63, 0.73]	0.81 [0.64, 0.93]
Rodrigues 2008	76	1	35	7	Occlusal	0.68 [0.59, 0.77]	0.88 [0.47, 1.00]
Diniz 2019	45	8	21	14	Occlusal	0.68 [0.56, 0.79]	0.64 [0.41, 0.83]
Diniz 2019	45	4	21	18	Occlusal	0.68 [0.56, 0.79]	0.82 [0.60, 0.95]
Huth 2010	52	12	25	28	Occlusal	0.68 [0.56, 0.78]	0.70 [0.53, 0.83]
Bussaneli 2015	44	4	22	24	Occlusal	0.67 [0.54, 0.78]	0.86 [0.67, 0.96]
Kucukyilmaz 2015	107	1	57	35	Occlusal	0.65 [0.57, 0.72]	0.97 [0.85, 1.00]
Diniz 2011	30	3	17	3	Occlusal	0.64 [0.49, 0.77]	0.50 [0.12, 0.88]
Souza 2013	38	5	22	14	Occlusal	0.63 [0.50, 0.75]	0.74 [0.49, 0.91]
Rodrigues 2011	50	4	30	13	Occlusal	0.63 [0.51, 0.73]	0.76 [0.50, 0.93]
Mendes 2006	51	2	32	25	Occlusal	0.61 [0.50, 0.72]	0.93 [0.76, 0.99]
Bittar 2012	22	3	14	19	Occlusal	0.61 [0.43, 0.77]	0.86 [0.65, 0.97]
Rocha 2003	35	4	23	38	Occlusal	0.60 [0.47, 0.73]	0.90 [0.77, 0.97]
Aktan 2012	42	12	33	42	Occlusal	0.56 [0.44, 0.67]	0.78 [0.64, 0.88]
Cinar 2013	18	0	15	11	Occlusal	0.55 [0.36, 0.72]	1.00 [0.72, 1.00]
Rodrigues 2009	83	1	74	11	Occlusal	0.53 [0.45, 0.61]	0.92 [0.62, 1.00]
Sheehy 2001	49	0	44	77	Occlusal	0.53 [0.42, 0.63]	1.00 [0.95, 1.00]
Van Hilsen 2013	16	2	16	8	Occlusal	0.50 [0.32, 0.68]	0.80 [0.44, 0.97]
Shi 2000	23	1	28	18	Occlusal	0.45 [0.31, 0.60]	0.95 [0.74, 1.00]
Matos 2011	155	3	195	29	Occlusal	0.44 [0.39, 0.50]	0.91 [0.75, 0.98]
Rodrigues 2011	25	1	55	16	Occlusal	0.31 [0.21, 0.43]	0.94 [0.71, 1.00]
Rodrigues 2009	33	1	103	11	Occlusal	0.24 [0.17, 0.32]	0.92 [0.62, 1.00]
Novaes 2016	73	26	3	7	Smooth	0.96 [0.89, 0.99]	0.21 [0.09, 0.39]
Novaes 2016	68	18	8	15	Smooth	0.89 [0.80, 0.95]	0.45 [0.28, 0.64]
Novaes 2016	59	6	17	27	Smooth	0.78 [0.67, 0.86]	0.82 [0.65, 0.93]
Pinelli 2002	78	30	31	81	Smooth	0.72 [0.62, 0.80]	0.73 [0.64, 0.81]
Almosa 2014	317	34	223	1079	Smooth	0.59 [0.54, 0.63]	0.97 [0.96, 0.98]
Mendes 2005	27	1	26	23	Smooth	0.51 [0.37, 0.65]	0.96 [0.79, 1.00]

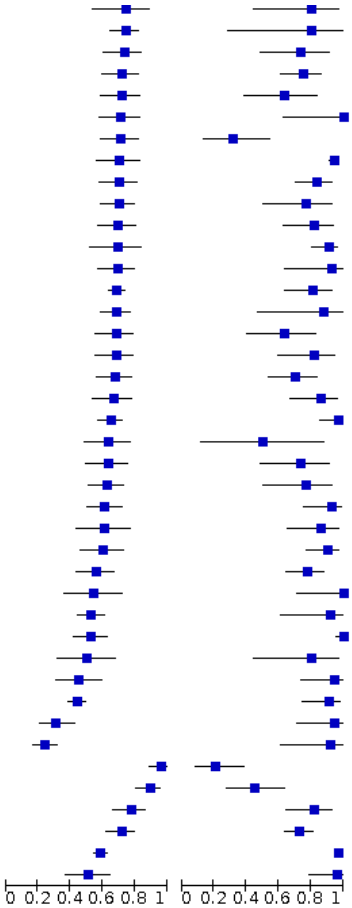
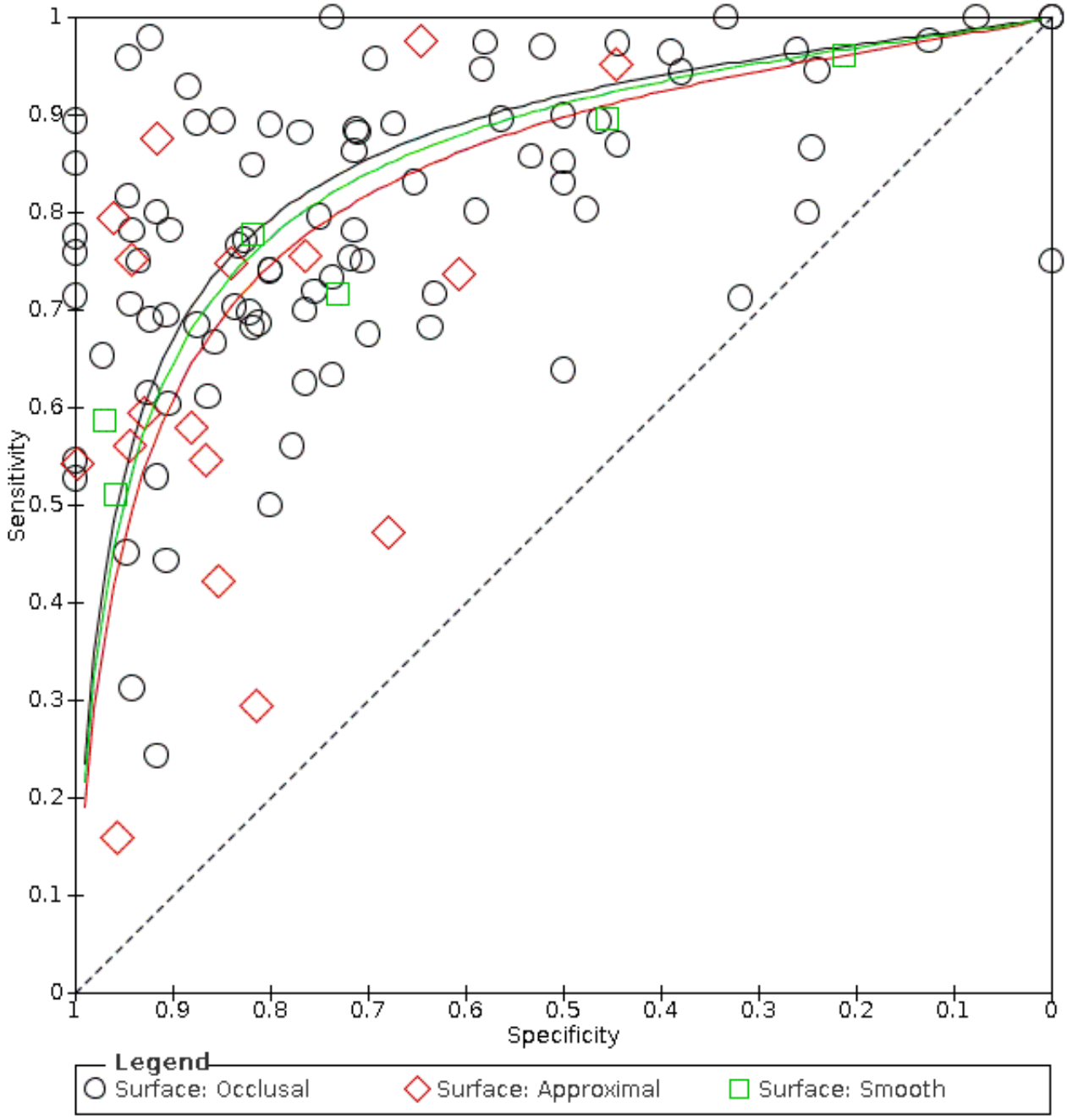


Figure 15. Summary receiver operating characteristic (SROC) plot presented according to tooth surface (proximal n = 18, occlusal n = 89, smooth/secondary caries n = 7).



Reference standard

The reference standard was either histology (78 datasets), enhanced visual assessment (25 datasets: Almosa 2014; Bussaneli 2015a; Chong 2003; Duruturk 2011; Feng 2005; Jablonski-Momeni 2014; Jablonski-Momeni 2016; Jung 2018; Kouchaji 2012; Kuhnisch 2008; Mansour 2016; Matos 2011; Mortensen 2018; Muller-Bolla 2017; Novaes 2009; Novaes 2010; Pinelli 2002; Presoto 2017; Sheehy 2001; Souza 2018; Umemori 2010; Zeitouny 2014), radiograph (six datasets: Kim 2017; Mepparambath 2014; Ribeiro 2015; Yoon 2017), or excavation (five datasets: Akarsu 2006; Bahrololoomi 2015; Chen 2012; Heinrich-Weltzien 2003; Huth 2010). The forest plots have been displayed arranged according to the reference standard (Figure 16) and results for the different reference standards were plotted in ROC space, with the HSROC curve plotted for each category (Figure 17). For the purpose of analysis excavation and histology were combined. When covariates for reference standard were included, removing shape from the model resulted in a negligible change in estimates (Chi^2 = 2.19, df = 2, P = 0.33). Whilst

there was some indication of a difference in curves according to the reference standard, when the models were tested for a difference in accuracy while leaving the shape of the curve consistent there was no statistical evidence of a difference in diagnostic accuracy across the groups (Chi^2 = 5.69, df = 2, P = 0.06). The RDOR for radiographs was 0.46 (95% CI 0.18 to 1.16), and for enhanced visual examination was 1.43 (95% CI 0.85 to 2.41) (Additional Table 4) when compared with the reference category of histology or excavation.

Figure 16. Forest plot of datasets categorised by reference standard (excavation n = 5, histology n = 78, radiograph n = 6, visual n = 25) and ordered by sensitivity.

Study	TP	FP	FN	TN	Reference standard	Sensitivity (95% CI)	Specificity (95% CI)
Heinrich-Weltzien 2003	212	10	12	14	Excavation	0.95 [0.91, 0.97]	0.58 [0.37, 0.78]
Akarsu 2006	112	11	15	27	Excavation	0.88 [0.81, 0.93]	0.71 [0.54, 0.85]
Bahrololoomi 2015	88	2	14	5	Excavation	0.86 [0.78, 0.92]	0.71 [0.29, 0.96]
Huth 2010	52	12	25	28	Excavation	0.68 [0.56, 0.78]	0.70 [0.53, 0.83]
Chen 2012	72	7	56	121	Excavation	0.56 [0.47, 0.65]	0.95 [0.89, 0.98]
Achilleos 2013	36	2	0	0	Histology	1.00 [0.90, 1.00]	0.00 [0.00, 0.84]
Baseren 2003	12	5	0	14	Histology	1.00 [0.74, 1.00]	0.74 [0.49, 0.91]
Jablonski-Momeni 2012a	67	12	0	1	Histology	1.00 [0.95, 1.00]	0.08 [0.00, 0.36]
Sridhar 2009	44	4	0	2	Histology	1.00 [0.92, 1.00]	0.33 [0.04, 0.78]
Kockanat 2017	92	2	2	24	Histology	0.98 [0.93, 1.00]	0.92 [0.75, 0.99]
Seremidi 2012	81	21	2	3	Histology	0.98 [0.92, 1.00]	0.13 [0.03, 0.32]
Bittar 2012	36	10	1	8	Histology	0.97 [0.86, 1.00]	0.44 [0.22, 0.69]
Jablonski-Momeni 2012	57	17	2	6	Histology	0.97 [0.88, 1.00]	0.26 [0.10, 0.48]
Pereira 2011	53	25	2	16	Histology	0.96 [0.87, 1.00]	0.39 [0.24, 0.55]
Novaes 2016	73	26	3	7	Histology	0.96 [0.89, 0.99]	0.21 [0.09, 0.39]
Lussi 2006a	89	8	4	18	Histology	0.96 [0.89, 0.99]	0.69 [0.48, 0.86]
Jablonski-Momeni 2011	69	19	4	6	Histology	0.95 [0.87, 0.98]	0.24 [0.09, 0.45]
Apostolopoulou 2009	98	1	11	1	Histology	0.90 [0.83, 0.95]	0.50 [0.01, 0.99]
Novaes 2016	68	18	8	15	Histology	0.89 [0.80, 0.95]	0.45 [0.28, 0.64]
Kockanat 2017	84	0	10	26	Histology	0.89 [0.81, 0.95]	1.00 [0.87, 1.00]
Aktan 2012	67	29	8	25	Histology	0.89 [0.80, 0.95]	0.46 [0.33, 0.60]
Rodrigues 2008	99	1	12	7	Histology	0.89 [0.82, 0.94]	0.88 [0.47, 1.00]
Diniz 2012	89	1	11	4	Histology	0.89 [0.81, 0.94]	0.80 [0.28, 0.99]
Lussi 2006a	82	6	11	20	Histology	0.88 [0.80, 0.94]	0.77 [0.56, 0.91]
Lussi 2006	78	5	11	56	Histology	0.88 [0.79, 0.94]	0.92 [0.82, 0.97]
Teo 2014	40	10	6	8	Histology	0.87 [0.74, 0.95]	0.44 [0.22, 0.69]
Braga 2009	71	37	11	12	Histology	0.87 [0.77, 0.93]	0.24 [0.13, 0.39]
Iranzo-Cortes 2017	42	7	7	8	Histology	0.86 [0.73, 0.94]	0.53 [0.27, 0.79]
Goel 2009	69	1	12	1	Histology	0.85 [0.76, 0.92]	0.50 [0.01, 0.99]
Diniz 2012	85	0	15	5	Histology	0.85 [0.76, 0.91]	1.00 [0.48, 1.00]
Cinar 2013	28	2	5	9	Histology	0.85 [0.68, 0.95]	0.82 [0.48, 0.98]
Jablonski-Momeni 2012	49	8	10	15	Histology	0.83 [0.71, 0.92]	0.65 [0.43, 0.84]
Diniz 2011	39	3	8	3	Histology	0.83 [0.69, 0.92]	0.50 [0.12, 0.88]
Diniz 2019	53	11	13	10	Histology	0.80 [0.69, 0.89]	0.48 [0.26, 0.70]
Castilho 2016	28	6	7	2	Histology	0.80 [0.63, 0.92]	0.25 [0.03, 0.65]
Lee 2018	40	1	10	11	Histology	0.80 [0.66, 0.90]	0.92 [0.62, 1.00]
Souza 2014	39	2	10	51	Histology	0.80 [0.66, 0.90]	0.96 [0.87, 1.00]
Seremidi 2012	66	6	17	18	Histology	0.80 [0.69, 0.88]	0.75 [0.53, 0.90]
Pereira 2011	43	4	12	37	Histology	0.78 [0.65, 0.88]	0.90 [0.77, 0.97]
Costa 2002	25	1	7	16	Histology	0.78 [0.60, 0.91]	0.94 [0.71, 1.00]
Novaes 2012a	50	14	14	35	Histology	0.78 [0.66, 0.87]	0.71 [0.57, 0.83]
Novaes 2016	59	6	17	27	Histology	0.78 [0.67, 0.86]	0.82 [0.65, 0.93]
Rodrigues 2008	86	0	25	8	Histology	0.77 [0.69, 0.85]	1.00 [0.63, 1.00]
Attrill 2001	27	4	8	19	Histology	0.77 [0.60, 0.90]	0.83 [0.61, 0.95]
Diniz 2011	36	1	11	5	Histology	0.77 [0.62, 0.88]	0.83 [0.36, 1.00]
Bamzahim 2004	26	6	8	26	Histology	0.76 [0.59, 0.89]	0.81 [0.64, 0.93]
Virajsilp 2005	67	1	22	17	Histology	0.75 [0.65, 0.84]	0.94 [0.73, 1.00]
Achilleos 2013	27	2	9	0	Histology	0.75 [0.58, 0.88]	0.00 [0.00, 0.84]
Ko 2015	57	3	19	16	Histology	0.75 [0.64, 0.84]	0.84 [0.60, 0.97]
Rodrigues 2011	60	5	20	12	Histology	0.75 [0.64, 0.84]	0.71 [0.44, 0.90]
Neuhaus 2011	20	2	7	8	Histology	0.74 [0.54, 0.89]	0.80 [0.44, 0.97]
Diniz 2012	74	1	26	4	Histology	0.74 [0.64, 0.82]	0.80 [0.28, 0.99]
Souza 2013	44	5	16	14	Histology	0.73 [0.60, 0.84]	0.74 [0.49, 0.91]
Novaes 2012a	46	12	18	37	Histology	0.72 [0.59, 0.82]	0.76 [0.61, 0.87]
Souza 2013	43	7	17	12	Histology	0.72 [0.59, 0.83]	0.63 [0.38, 0.84]
Paula 2011	40	0	16	8	Histology	0.71 [0.58, 0.83]	1.00 [0.63, 1.00]
Diniz 2019	47	15	19	7	Histology	0.71 [0.59, 0.82]	0.32 [0.14, 0.55]
Novaes 2012a	45	8	19	41	Histology	0.70 [0.58, 0.81]	0.84 [0.70, 0.93]
Rodrigues 2011	56	4	24	13	Histology	0.70 [0.59, 0.80]	0.76 [0.50, 0.93]
Bussaneli 2015	46	5	20	23	Histology	0.70 [0.57, 0.80]	0.82 [0.63, 0.94]
Jablonski-Momeni 2012a	49	1	22	12	Histology	0.69 [0.57, 0.79]	0.92 [0.64, 1.00]
Rodrigues 2008	76	1	35	7	Histology	0.68 [0.59, 0.77]	0.88 [0.47, 1.00]
Diniz 2019	45	8	21	14	Histology	0.68 [0.56, 0.79]	0.64 [0.41, 0.83]
Diniz 2019	45	4	21	18	Histology	0.68 [0.56, 0.79]	0.82 [0.60, 0.95]
Bussaneli 2015	44	4	22	24	Histology	0.67 [0.54, 0.78]	0.86 [0.67, 0.96]
Kucukyilmaz 2015	107	1	57	35	Histology	0.65 [0.57, 0.72]	0.97 [0.85, 1.00]
Diniz 2011	30	3	17	3	Histology	0.64 [0.49, 0.77]	0.50 [0.12, 0.88]
Souza 2013	38	5	22	14	Histology	0.63 [0.50, 0.75]	0.74 [0.49, 0.91]
Rodrigues 2011	50	4	30	13	Histology	0.63 [0.51, 0.73]	0.76 [0.50, 0.93]
Mendes 2006	51	2	32	25	Histology	0.61 [0.50, 0.72]	0.93 [0.76, 0.99]
Bittar 2012	22	3	14	19	Histology	0.61 [0.43, 0.77]	0.86 [0.65, 0.97]
Rocha 2003	35	4	23	38	Histology	0.60 [0.47, 0.73]	0.90 [0.77, 0.97]
Ozsevik 2015	59	4	40	53	Histology	0.60 [0.49, 0.69]	0.93 [0.83, 0.98]
Souza 2014	29	11	21	83	Histology	0.58 [0.43, 0.72]	0.88 [0.80, 0.94]
Aktan 2012	42	12	33	42	Histology	0.56 [0.44, 0.67]	0.78 [0.64, 0.88]
Cinar 2013	18	0	15	11	Histology	0.55 [0.36, 0.72]	1.00 [0.72, 1.00]
Tonkaboni 2018	25	0	21	62	Histology	0.54 [0.39, 0.69]	1.00 [0.94, 1.00]
Rodrigues 2009	83	1	74	11	Histology	0.53 [0.45, 0.61]	0.92 [0.62, 1.00]
Mendes 2005	27	1	26	23	Histology	0.51 [0.37, 0.65]	0.96 [0.79, 1.00]
Van Hilsen 2013	16	2	16	8	Histology	0.50 [0.32, 0.68]	0.80 [0.44, 0.97]
Shi 2000	23	1	28	18	Histology	0.45 [0.31, 0.60]	0.95 [0.74, 1.00]

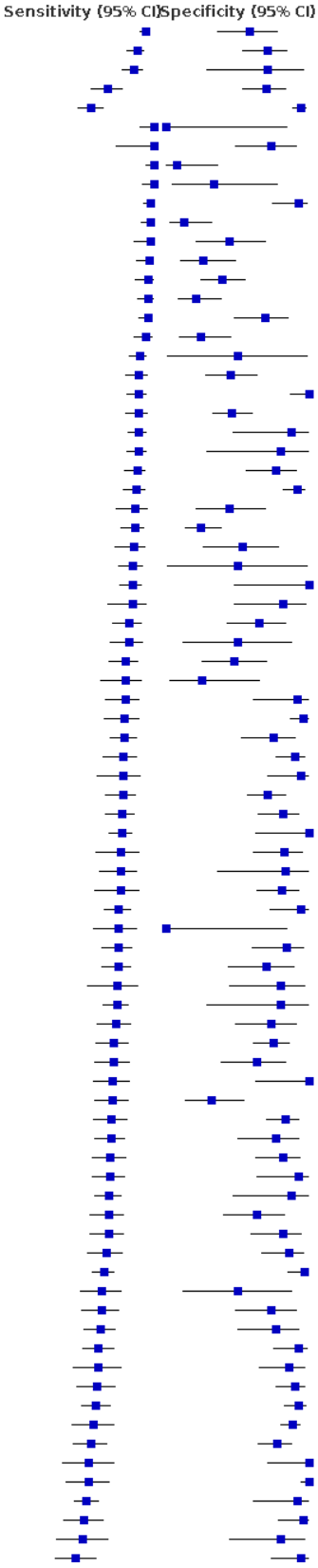


Figure 16. (Continued)

Mendes 2005	27	1	26	23	Histology	0.51 [0.37, 0.65]	0.96 [0.79, 1.00]
Van Hilsen 2013	16	2	16	8	Histology	0.50 [0.32, 0.68]	0.80 [0.44, 0.97]
Shi 2000	23	1	28	18	Histology	0.45 [0.31, 0.60]	0.95 [0.74, 1.00]
Rodrigues 2011	25	1	55	16	Histology	0.31 [0.21, 0.43]	0.94 [0.71, 1.00]
Rodrigues 2009	33	1	103	11	Histology	0.24 [0.17, 0.32]	0.92 [0.62, 1.00]
Yoon 2017	61	21	3	17	Radiograph	0.95 [0.87, 0.99]	0.45 [0.29, 0.62]
Rando-Meirelles 2011	213	215	53	308	Radiograph	0.80 [0.75, 0.85]	0.59 [0.55, 0.63]
Mepparambath 2014	28	31	9	101	Radiograph	0.76 [0.59, 0.88]	0.77 [0.68, 0.83]
Yoon 2017	35	5	29	33	Radiograph	0.55 [0.42, 0.67]	0.87 [0.72, 0.96]
Ribeiro 2015	18	8	20	17	Radiograph	0.47 [0.31, 0.64]	0.68 [0.46, 0.85]
Kim 2017	72	16	98	94	Radiograph	0.42 [0.35, 0.50]	0.85 [0.77, 0.91]
Chong 2003	48	272	0	0	Visual	1.00 [0.93, 1.00]	0.00 [0.00, 0.01]
Bussaneli 2015a	41	6	1	11	Visual	0.98 [0.87, 1.00]	0.65 [0.38, 0.86]
Presoto 2017	37	29	1	40	Visual	0.97 [0.86, 1.00]	0.58 [0.45, 0.70]
Kouchaji 2012	129	11	4	12	Visual	0.97 [0.92, 0.99]	0.52 [0.31, 0.73]
Feng 2005	342	75	15	1300	Visual	0.96 [0.93, 0.98]	0.95 [0.93, 0.96]
Jablonski-Momeni 2014	50	157	3	96	Visual	0.94 [0.84, 0.99]	0.38 [0.32, 0.44]
Zeitouny 2014	104	6	8	46	Visual	0.93 [0.86, 0.97]	0.88 [0.77, 0.96]
Jung 2018	333	183	39	236	Visual	0.90 [0.86, 0.92]	0.56 [0.51, 0.61]
Jablonski-Momeni 2016	67	21	8	118	Visual	0.89 [0.80, 0.95]	0.85 [0.78, 0.90]
Duruturk 2011	163	105	20	217	Visual	0.89 [0.84, 0.93]	0.67 [0.62, 0.72]
Muller-Bolla 2017	473	60	62	148	Visual	0.88 [0.85, 0.91]	0.71 [0.64, 0.77]
Kuhnisch 2008	489	13	110	228	Visual	0.82 [0.78, 0.85]	0.95 [0.91, 0.97]
Mortensen 2018	44	0	14	2	Visual	0.76 [0.63, 0.86]	1.00 [0.16, 1.00]
Muller-Bolla 2017	404	58	133	148	Visual	0.75 [0.71, 0.79]	0.72 [0.65, 0.78]
Jablonski-Momeni 2016	27	11	9	158	Visual	0.75 [0.58, 0.88]	0.93 [0.89, 0.97]
Souza 2018	59	45	21	70	Visual	0.74 [0.63, 0.83]	0.61 [0.51, 0.70]
Pinelli 2002	78	30	31	81	Visual	0.72 [0.62, 0.80]	0.73 [0.64, 0.81]
Mansour 2016	36	21	15	354	Visual	0.71 [0.56, 0.83]	0.94 [0.92, 0.97]
Umemori 2010	25	6	11	58	Visual	0.69 [0.52, 0.84]	0.91 [0.81, 0.96]
Matos 2011	240	6	110	26	Visual	0.69 [0.63, 0.73]	0.81 [0.64, 0.93]
Almosa 2014	317	34	223	1079	Visual	0.59 [0.54, 0.63]	0.97 [0.96, 0.98]
Sheehy 2001	49	0	44	77	Visual	0.53 [0.42, 0.63]	1.00 [0.95, 1.00]
Matos 2011	155	3	195	29	Visual	0.44 [0.39, 0.50]	0.91 [0.75, 0.98]
Novaes 2010	141	21	337	93	Visual	0.29 [0.25, 0.34]	0.82 [0.73, 0.88]
Novaes 2009	41	15	215	350	Visual	0.16 [0.12, 0.21]	0.96 [0.93, 0.98]

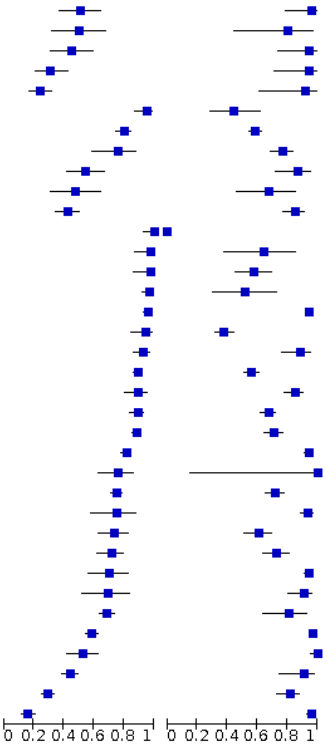
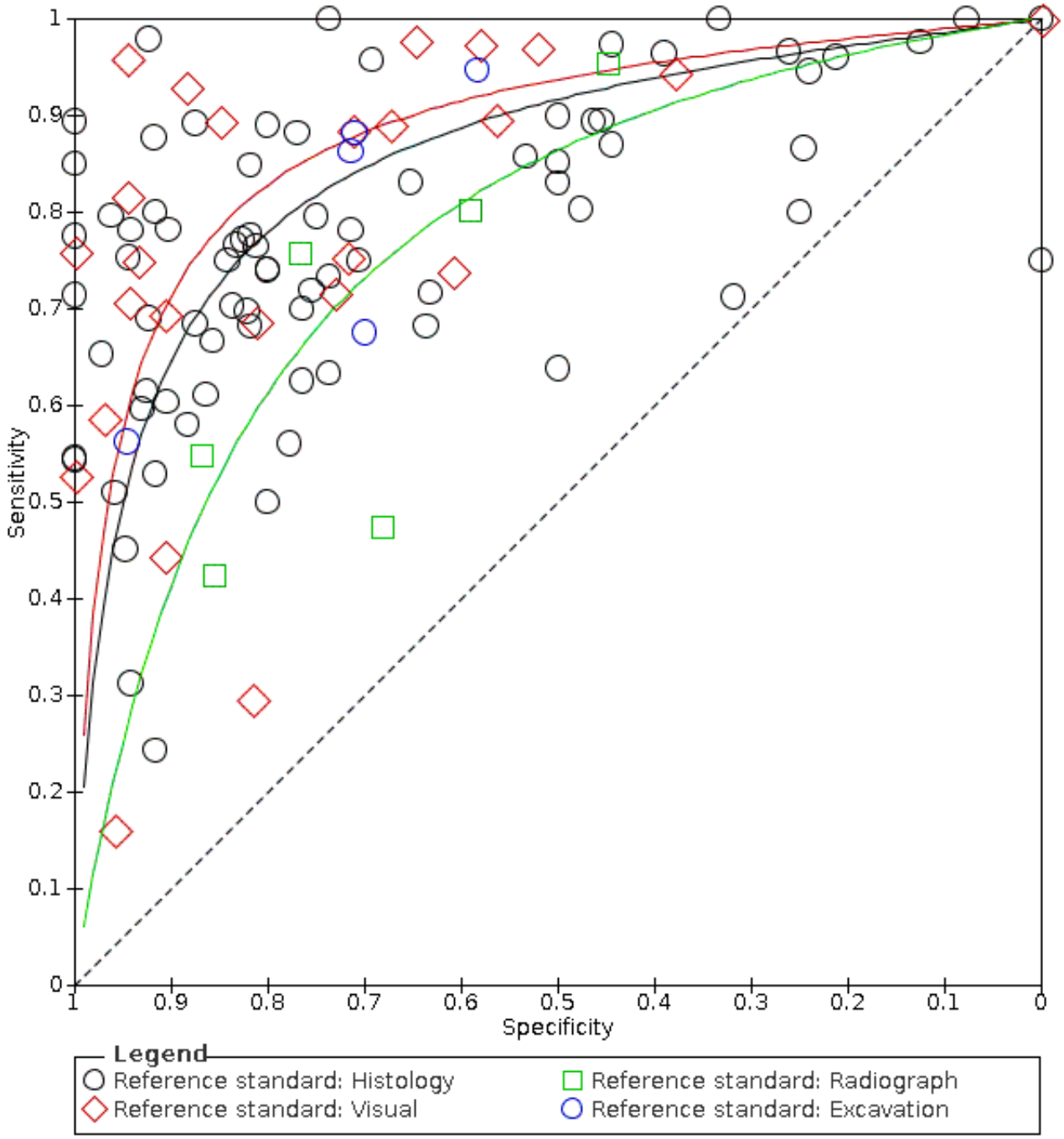


Figure 17. Summary receiver operating characteristic (SROC) plot presented according to reference standard (histology or excavation n = 83, radiograph n = 6, enhanced visual examination n = 25).



Multiple sites

We planned to investigate the effect of assessments at the patient, tooth, or at multiple sites per tooth level. No studies reported at the patient level but there were 24 datasets that reported multiple sites per tooth and the remaining 90 datasets reported one site per tooth. The forest plots are sorted to show those with multiple sites first, these are then arranged by sensitivity (Figure 18). The two groups were plotted in ROC space and SROC curves plotted for each group (Figure 19).

Figure 18. Forest plot of all studies investigating the effect of multiple sites per tooth.

Study	TP	FP	FN	TN	Multiple sites	Sensitivity (95% CI)	Specificity (95% CI)
Seremidi 2012	81	21	2	3	Yes	0.98 [0.92, 1.00]	0.13 [0.03, 0.32]
Jablonski-Momeni 2012	57	17	2	6	Yes	0.97 [0.88, 1.00]	0.26 [0.10, 0.48]
Novaes 2016	73	26	3	7	Yes	0.96 [0.89, 0.99]	0.21 [0.09, 0.39]
Jablonski-Momeni 2011	69	19	4	6	Yes	0.95 [0.87, 0.98]	0.24 [0.09, 0.45]
Apostolopoulou 2009	98	1	11	1	Yes	0.90 [0.83, 0.95]	0.50 [0.01, 0.99]
Novaes 2016	68	18	8	15	Yes	0.89 [0.80, 0.95]	0.45 [0.28, 0.64]
Aktan 2012	67	29	8	25	Yes	0.89 [0.80, 0.95]	0.46 [0.33, 0.60]
Duruturk 2011	163	105	20	217	Yes	0.89 [0.84, 0.93]	0.67 [0.62, 0.72]
Lussi 2006	78	5	11	56	Yes	0.88 [0.79, 0.94]	0.92 [0.82, 0.97]
Jablonski-Momeni 2012	49	8	10	15	Yes	0.83 [0.71, 0.92]	0.65 [0.43, 0.84]
Seremidi 2012	66	6	17	18	Yes	0.80 [0.69, 0.88]	0.75 [0.53, 0.90]
Novaes 2012a	50	14	14	35	Yes	0.78 [0.66, 0.87]	0.71 [0.57, 0.83]
Novaes 2016	59	6	17	27	Yes	0.78 [0.67, 0.86]	0.82 [0.65, 0.93]
Novaes 2012a	46	12	18	37	Yes	0.72 [0.59, 0.82]	0.76 [0.61, 0.87]
Novaes 2012a	45	8	19	41	Yes	0.70 [0.58, 0.81]	0.84 [0.70, 0.93]
Matos 2011	240	6	110	26	Yes	0.69 [0.63, 0.73]	0.81 [0.64, 0.93]
Mendes 2006	51	2	32	25	Yes	0.61 [0.50, 0.72]	0.93 [0.76, 0.99]
Rocha 2003	35	4	23	38	Yes	0.60 [0.47, 0.73]	0.90 [0.77, 0.97]
Ozsevik 2015	59	4	40	53	Yes	0.60 [0.49, 0.69]	0.93 [0.83, 0.98]
Chen 2012	72	7	56	121	Yes	0.56 [0.47, 0.65]	0.95 [0.89, 0.98]
Aktan 2012	42	12	33	42	Yes	0.56 [0.44, 0.67]	0.78 [0.64, 0.88]
Ribeiro 2015	18	8	20	17	Yes	0.47 [0.31, 0.64]	0.68 [0.46, 0.85]
Matos 2011	155	3	195	29	Yes	0.44 [0.39, 0.50]	0.91 [0.75, 0.98]
Novaes 2010	141	21	337	93	Yes	0.29 [0.25, 0.34]	0.82 [0.73, 0.88]
Achilleos 2013	36	2	0	0	No	1.00 [0.90, 1.00]	0.00 [0.00, 0.84]
Baseren 2003	12	5	0	14	No	1.00 [0.74, 1.00]	0.74 [0.49, 0.91]
Chong 2003	48	272	0	0	No	1.00 [0.93, 1.00]	0.00 [0.00, 0.01]
Jablonski-Momeni 2012a	67	12	0	1	No	1.00 [0.95, 1.00]	0.08 [0.00, 0.36]
Sridhar 2009	44	4	0	2	No	1.00 [0.92, 1.00]	0.33 [0.04, 0.78]
Kockanat 2017	92	2	2	24	No	0.98 [0.93, 1.00]	0.92 [0.75, 0.99]
Bussaneli 2015a	41	6	1	11	No	0.98 [0.87, 1.00]	0.65 [0.38, 0.86]
Presoto 2017	37	29	1	40	No	0.97 [0.86, 1.00]	0.58 [0.45, 0.70]
Bittar 2012	36	10	1	8	No	0.97 [0.86, 1.00]	0.44 [0.22, 0.69]
Kouchaji 2012	129	11	4	12	No	0.97 [0.92, 0.99]	0.52 [0.31, 0.73]
Pereira 2011	53	25	2	16	No	0.96 [0.87, 1.00]	0.39 [0.24, 0.55]
Feng 2005	342	75	15	1300	No	0.96 [0.93, 0.98]	0.95 [0.93, 0.96]
Lussi 2006a	89	8	4	18	No	0.96 [0.89, 0.99]	0.69 [0.48, 0.86]
Yoon 2017	61	21	3	17	No	0.95 [0.87, 0.99]	0.45 [0.29, 0.62]
Heinrich-Weltzien 2003	212	10	12	14	No	0.95 [0.91, 0.97]	0.58 [0.37, 0.78]
Jablonski-Momeni 2014	50	157	3	96	No	0.94 [0.84, 0.99]	0.38 [0.32, 0.44]
Zeitouny 2014	104	6	8	46	No	0.93 [0.86, 0.97]	0.88 [0.77, 0.96]
Jung 2018	333	183	39	236	No	0.90 [0.86, 0.92]	0.56 [0.51, 0.61]
Kockanat 2017	84	0	10	26	No	0.89 [0.81, 0.95]	1.00 [0.87, 1.00]
Jablonski-Momeni 2016	67	21	8	118	No	0.89 [0.80, 0.95]	0.85 [0.78, 0.90]
Rodrigues 2008	99	1	12	7	No	0.89 [0.82, 0.94]	0.88 [0.47, 1.00]
Diniz 2012	89	1	11	4	No	0.89 [0.81, 0.94]	0.80 [0.28, 0.99]
Muller-Bolla 2017	473	60	62	148	No	0.88 [0.85, 0.91]	0.71 [0.64, 0.77]
Akarsu 2006	112	11	15	27	No	0.88 [0.81, 0.93]	0.71 [0.54, 0.85]
Lussi 2006a	82	6	11	20	No	0.88 [0.80, 0.94]	0.77 [0.56, 0.91]
Teo 2014	40	10	6	8	No	0.87 [0.74, 0.95]	0.44 [0.22, 0.69]
Braga 2009	71	37	11	12	No	0.87 [0.77, 0.93]	0.24 [0.13, 0.39]
Bahrololoomi 2015	88	2	14	5	No	0.86 [0.78, 0.92]	0.71 [0.29, 0.96]
Iranzo-Cortes 2017	42	7	7	8	No	0.86 [0.73, 0.94]	0.53 [0.27, 0.79]
Goel 2009	69	1	12	1	No	0.85 [0.76, 0.92]	0.50 [0.01, 0.99]
Diniz 2012	85	0	15	5	No	0.85 [0.76, 0.91]	1.00 [0.48, 1.00]
Cinar 2013	28	2	5	9	No	0.85 [0.68, 0.95]	0.82 [0.48, 0.98]
Diniz 2011	39	3	8	3	No	0.83 [0.69, 0.92]	0.50 [0.12, 0.88]
Kuhnisch 2008	489	13	110	228	No	0.82 [0.78, 0.85]	0.95 [0.91, 0.97]
Diniz 2019	53	11	13	10	No	0.80 [0.69, 0.89]	0.48 [0.26, 0.70]
Rando-Meirelles 2011	213	215	53	308	No	0.80 [0.75, 0.85]	0.59 [0.55, 0.63]
Castilho 2016	28	6	7	2	No	0.80 [0.63, 0.92]	0.25 [0.03, 0.65]
Lee 2018	40	1	10	11	No	0.80 [0.66, 0.90]	0.92 [0.62, 1.00]
Souza 2014	39	2	10	51	No	0.80 [0.66, 0.90]	0.96 [0.87, 1.00]
Pereira 2011	43	4	12	37	No	0.78 [0.65, 0.88]	0.90 [0.77, 0.97]
Costa 2002	25	1	7	16	No	0.78 [0.60, 0.91]	0.94 [0.71, 1.00]
Rodrigues 2008	86	0	25	8	No	0.77 [0.69, 0.85]	1.00 [0.63, 1.00]
Attrill 2001	27	4	8	19	No	0.77 [0.60, 0.90]	0.83 [0.61, 0.95]
Diniz 2011	36	1	11	5	No	0.77 [0.62, 0.88]	0.83 [0.36, 1.00]
Bamzahim 2004	26	6	8	26	No	0.76 [0.59, 0.89]	0.81 [0.64, 0.93]
Mortensen 2018	44	0	14	2	No	0.76 [0.63, 0.86]	1.00 [0.16, 1.00]
Mepparambath 2014	28	31	9	101	No	0.76 [0.59, 0.88]	0.77 [0.68, 0.83]
Virajsilp 2005	67	1	22	17	No	0.75 [0.65, 0.84]	0.94 [0.73, 1.00]
Muller-Bolla 2017	404	58	133	148	No	0.75 [0.71, 0.79]	0.72 [0.65, 0.78]
Achilleos 2013	27	2	9	0	No	0.75 [0.58, 0.88]	0.00 [0.00, 0.84]
Jablonski-Momeni 2016	27	11	9	158	No	0.75 [0.58, 0.88]	0.93 [0.89, 0.97]
Ko 2015	57	3	19	16	No	0.75 [0.64, 0.84]	0.84 [0.60, 0.97]
Rodrigues 2011	60	5	20	12	No	0.75 [0.64, 0.84]	0.71 [0.44, 0.90]

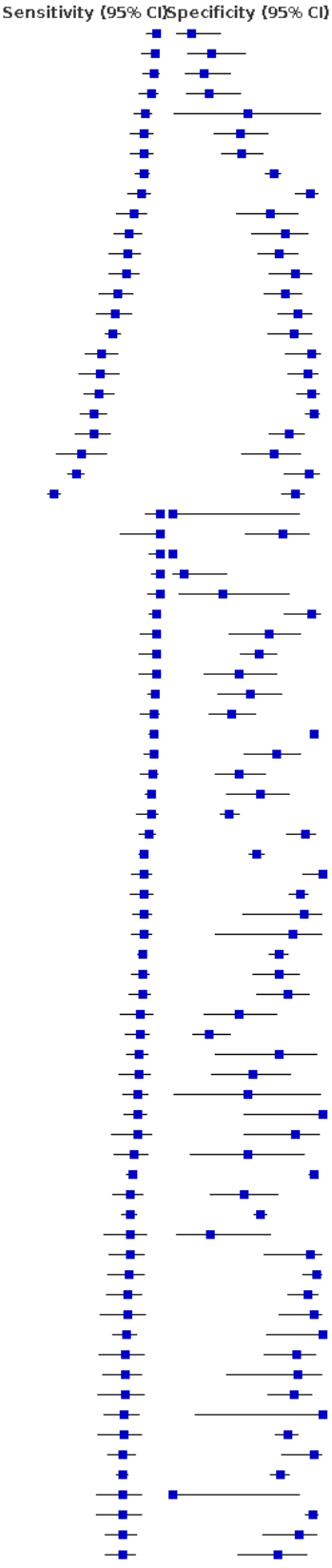


Figure 18. (Continued)

Jablonski-Momeni 2016	27	11	9	158	No	0.75 [0.58, 0.88]	0.93 [0.89, 0.97]
Ko 2015	57	3	19	16	No	0.75 [0.64, 0.84]	0.84 [0.60, 0.97]
Rodrigues 2011	60	5	20	12	No	0.75 [0.64, 0.84]	0.71 [0.44, 0.90]
Neuhaus 2011	20	2	7	8	No	0.74 [0.54, 0.89]	0.80 [0.44, 0.97]
Diniz 2012	74	1	26	4	No	0.74 [0.64, 0.82]	0.80 [0.28, 0.99]
Souza 2018	59	45	21	70	No	0.74 [0.63, 0.83]	0.61 [0.51, 0.70]
Souza 2013	44	5	16	14	No	0.73 [0.60, 0.84]	0.74 [0.49, 0.91]
Souza 2013	43	7	17	12	No	0.72 [0.59, 0.83]	0.63 [0.38, 0.84]
Pinelli 2002	78	30	31	81	No	0.72 [0.62, 0.80]	0.73 [0.64, 0.81]
Paula 2011	40	0	16	8	No	0.71 [0.58, 0.83]	1.00 [0.63, 1.00]
Diniz 2019	47	15	19	7	No	0.71 [0.59, 0.82]	0.32 [0.14, 0.55]
Mansour 2016	36	21	15	354	No	0.71 [0.56, 0.83]	0.94 [0.92, 0.97]
Rodrigues 2011	56	4	24	13	No	0.70 [0.59, 0.80]	0.76 [0.50, 0.93]
Bussaneli 2015	46	5	20	23	No	0.70 [0.57, 0.80]	0.82 [0.63, 0.94]
Umemori 2010	25	6	11	58	No	0.69 [0.52, 0.84]	0.91 [0.81, 0.96]
Jablonski-Momeni 2012a	49	1	22	12	No	0.69 [0.57, 0.79]	0.92 [0.64, 1.00]
Rodrigues 2008	76	1	35	7	No	0.68 [0.59, 0.77]	0.88 [0.47, 1.00]
Diniz 2019	45	8	21	14	No	0.68 [0.56, 0.79]	0.64 [0.41, 0.83]
Diniz 2019	45	4	21	18	No	0.68 [0.56, 0.79]	0.82 [0.60, 0.95]
Huth 2010	52	12	25	28	No	0.68 [0.56, 0.78]	0.70 [0.53, 0.83]
Bussaneli 2015	44	4	22	24	No	0.67 [0.54, 0.78]	0.86 [0.67, 0.96]
Kucukyilmaz 2015	107	1	57	35	No	0.65 [0.57, 0.72]	0.97 [0.85, 1.00]
Diniz 2011	30	3	17	3	No	0.64 [0.49, 0.77]	0.50 [0.12, 0.88]
Souza 2013	38	5	22	14	No	0.63 [0.50, 0.75]	0.74 [0.49, 0.91]
Rodrigues 2011	50	4	30	13	No	0.63 [0.51, 0.73]	0.76 [0.50, 0.93]
Bittar 2012	22	3	14	19	No	0.61 [0.43, 0.77]	0.86 [0.65, 0.97]
Almosa 2014	317	34	223	1079	No	0.59 [0.54, 0.63]	0.97 [0.96, 0.98]
Souza 2014	29	11	21	83	No	0.58 [0.43, 0.72]	0.88 [0.80, 0.94]
Yoon 2017	35	5	29	33	No	0.55 [0.42, 0.67]	0.87 [0.72, 0.96]
Cinar 2013	18	0	15	11	No	0.55 [0.36, 0.72]	1.00 [0.72, 1.00]
Tonkaboni 2018	25	0	21	62	No	0.54 [0.39, 0.69]	1.00 [0.94, 1.00]
Rodrigues 2009	83	1	74	11	No	0.53 [0.45, 0.61]	0.92 [0.62, 1.00]
Sheehy 2001	49	0	44	77	No	0.53 [0.42, 0.63]	1.00 [0.95, 1.00]
Mendes 2005	27	1	26	23	No	0.51 [0.37, 0.65]	0.96 [0.79, 1.00]
Van Hilsen 2013	16	2	16	8	No	0.50 [0.32, 0.68]	0.80 [0.44, 0.97]
Shi 2000	23	1	28	18	No	0.45 [0.31, 0.60]	0.95 [0.74, 1.00]
Kim 2017	72	16	98	94	No	0.42 [0.35, 0.50]	0.85 [0.77, 0.91]
Rodrigues 2011	25	1	55	16	No	0.31 [0.21, 0.43]	0.94 [0.71, 1.00]
Rodrigues 2009	33	1	103	11	No	0.24 [0.17, 0.32]	0.92 [0.62, 1.00]
Novaes 2009	41	15	215	350	No	0.16 [0.12, 0.21]	0.96 [0.93, 0.98]

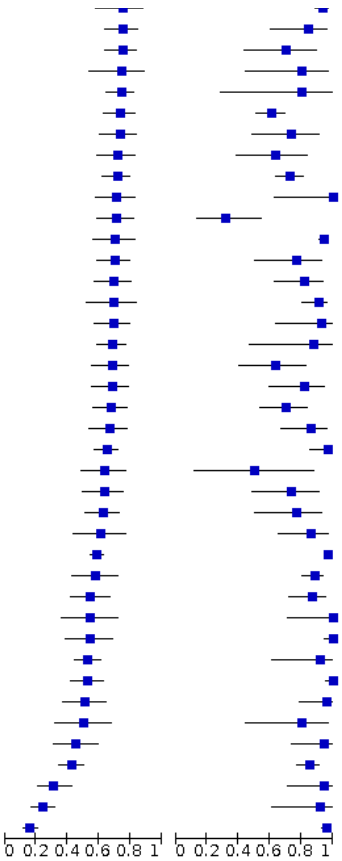
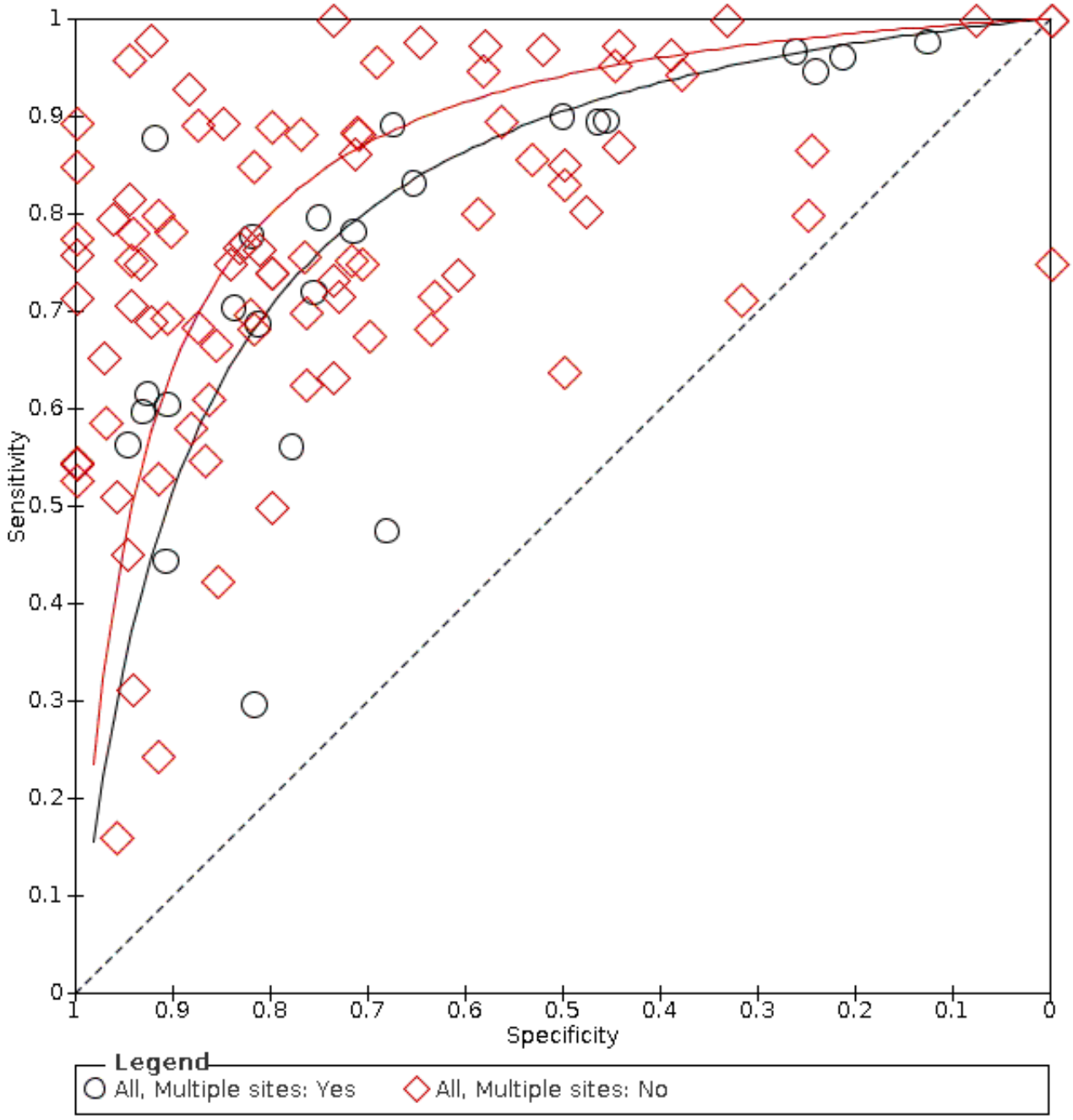


Figure 19. Summary receiver operating characteristic (SROC) plot presented according to multiple or single site (multiple sites per tooth n = 24 and single sites n = 90).



When covariates for the number of sites were included, removing shape from the model resulted in a negligible change in estimates (Chi^2 = 0.42, df = 2, P = 0.51). Whilst there was some indication of a difference in curves according to the number of sites, when the models were tested for a difference in accuracy while leaving the shape of the curve consistent across groups there was no statistical evidence of a difference of diagnostic accuracy between the groups (Chi^2 = 3.49, df = 1, P = 0.06). The RDOR for multiple sites was 0.59 (95% CI 0.35 to 1.02) (Additional Table 4) when compared with the reference category of single site assessment.

Sensitivity analysis

Sensitivity analysis was proposed a priori to investigate the effect of study quality on the sensitivity and specificity results. The highest proportion of high risk of bias assessments was observed in the participant selection domain (Figure 3) where only nine datasets (8%) in the meta-analysis were judged as at low risk of bias. Figure

20 shows the SROC plot with all included studies labelled according to low, unclear, or high risk of bias. Of the low risk of bias datasets, only two lie above the ROC curve (Almosa 2014; Zeitouny 2014). Figure 21 applies a sensitivity analysis and recalculates the ROC curve for the datasets which were allocated a low risk of bias for participant selection in QUADAS-2. This results in an ROC curve with lower sensitivity and specificity than the curve for all datasets. Formal statistical analysis was not performed due to the small number of datasets in the low risk of bias group.

Figure 20. Summary receiver operating characteristic (SROC) plot of all datasets with risk of bias for participant selection domain identified.

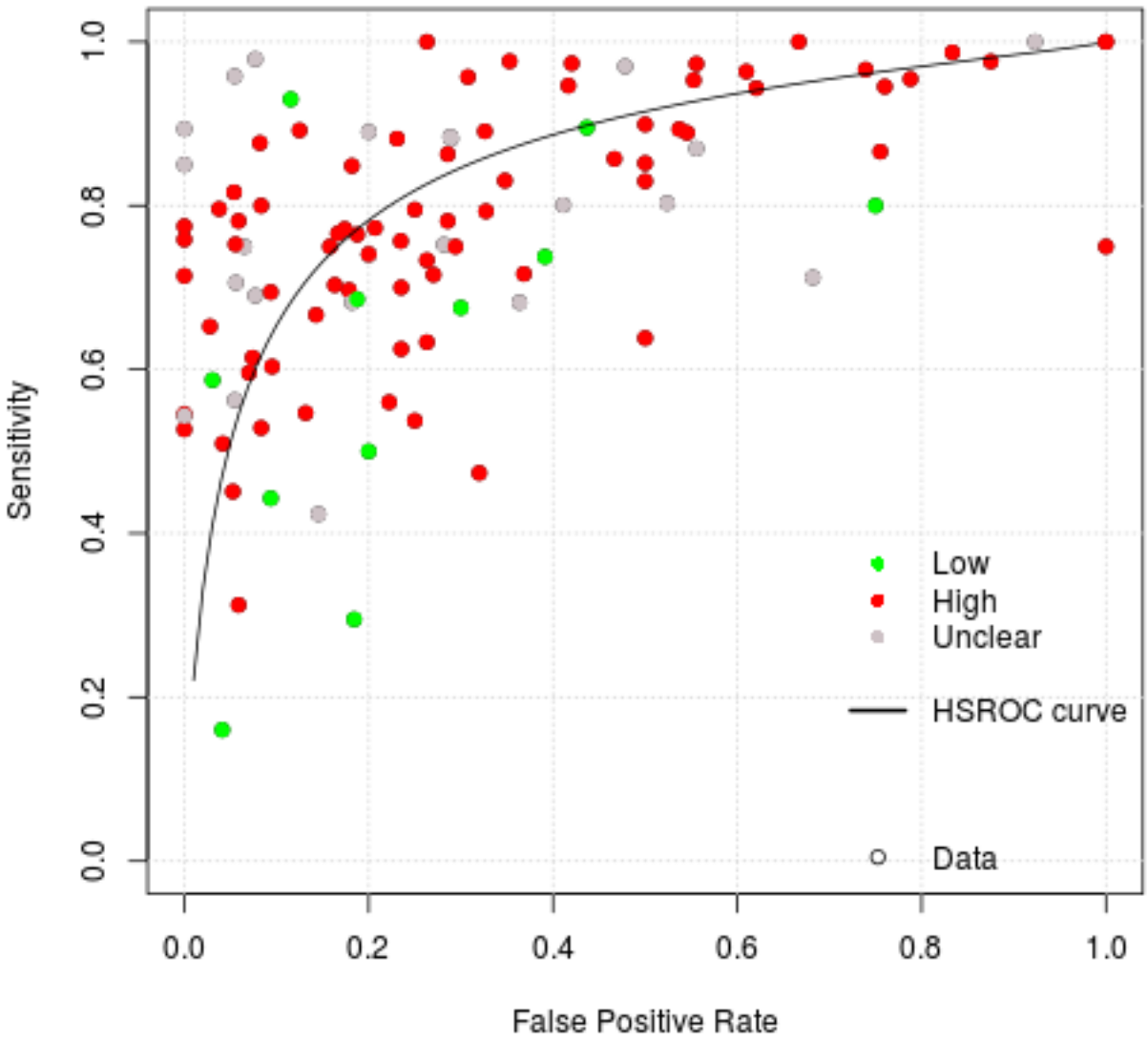
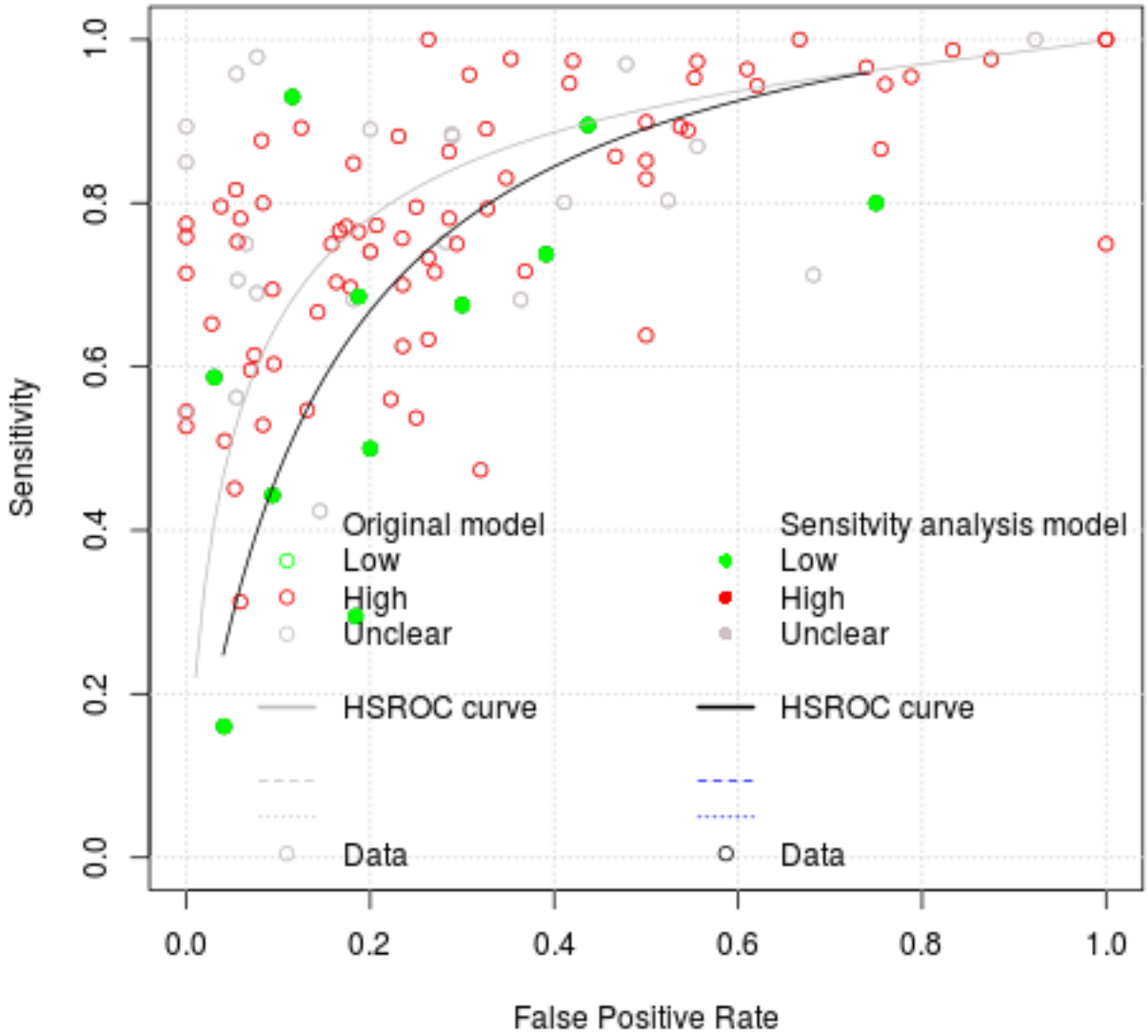


Figure 21. Sensitivity analysis of datasets reporting low risk of bias for participant selection domain.



DISCUSSION

Summary of main results

The included studies allowed us to evaluate the diagnostic test accuracy of fluorescence-based devices for the detection of early or non-cavitated caries, with particular focus on early-stage caries in the enamel of the tooth. A large number of studies were available that investigated fluorescence devices and they covered a range of different methods which utilise differing technologies, in particular by exploiting different wavelengths of light to perform the task of detecting caries.

Sufficient studies presented data in a format that allowed the construction of 2 x 2 tables and meta-analyses. However, there was substantial variation in values of sensitivity and specificity for each class of fluorescence devices and extensive heterogeneity in study design, sample population, index test, and reference standard. This is an important consideration for the interpretation of the results of this review. The low methodological quality of the available studies is partly due to unavoidable difficulties in study design, however, we judged one study as low risk across all domains of risk of bias and as low concern for applicability. Participant selection was the domain where we observed the highest percentage of high risk of bias judgements. The included patients, teeth, or surfaces should be recruited consecutively or randomly and the methods reported, thereby avoiding any suggestion that teeth are included that are more complex or straightforward to diagnose which would introduce bias. There were also substantial applicability concerns due to the inclusion of a large number of studies with an in vitro study design. Whilst we acknowledge that this is an important part of the development of diagnostic tests, these studies inevitably cause high concern for applicability to our research question which aimed to determine the accuracy of these devices in a clinical setting with the difficulties of access to the oral cavity, patient acceptability, and time constraints for examinations. The dominance of in vitro studies also means that the information on how the results of these devices are used to support diagnosis, as opposed to pure detection, is limited. In contrast to the participant

selection domain, the index test and reference standard domains showed a high number of studies with low risk of bias and applicability concern. Similarly, flow and timing were of concern in only 25% of the included studies. Reasons for high risk of bias judgements for the index test domain largely resulted from the lack of a pre-defined threshold. This was often because studies were attempting to determine the most appropriate threshold for their sample population, resulting in inflating sensitivity and specificity and therefore introducing bias. We awarded a decision of high risk where an imperfect reference standard, such as a visual examination or radiograph, was used. This highlights the main difficulty in studies of this type; to correctly classify the target condition, the preferred reference standard is histology. However, this automatically elicits concern for applicability in participant selection. The studies that circumvented this issue did so by targeting patients close to exfoliation of a primary tooth or those who required a tooth extraction and applied the index test in vivo with a subsequent reference standard in vitro. Such studies are challenging to organise and administer, and could still be considered to lack broader applicability since they often use teenage children requiring extractions for orthodontic purposes, and who would potentially have a lower prevalence of caries than adults.

We estimated the accuracy of any fluorescence-based device for the detection of early dental caries and compared the three groups of red, blue, and green fluorescence. These devices produced an outcome on a continuous scale and applied different thresholds to determine the result. Consequently we have used summary receiver operating characteristic (SROC) curves rather than summary sensitivity and specificity estimates. We took illustrative sensitivity values from the hierarchical summary receiver operating characteristic (HSROC) curves (at a fixed sensitivity of 0.78 (median) and 0.90 (upper quartile)) to illustrate changes in sensitivity and specificity along the HSROC curve. These values are intended to be used only as a guide and should not be used to indicate the actual performance of these fluorescence devices. We used meta-regression to explore potential sources of heterogeneity, but pre-specified patient or study characteristics were unable to account for the substantial variation in results.

One of the primary objectives of the review was to investigate the effect of using the fluorescence devices in combination with other tests, particularly as an adjunct to a visual examination. Only one study (Alomari 2015) formally reported this, and therefore it has not been possible to make an assessment. There were no case-control or randomised controlled trials included in this review, as the searches retrieved no such eligible studies. There were also a limited number of included studies that investigated the effect of sealants or restorations on the diagnostic accuracy of fluorescence tests.

Despite the relatively large volume of evidence we rated the certainty of the evidence as low, downgraded two levels in total, for risk of bias due to limitations in the design and conduct of the included studies, indirectness arising from the high number of in vitro studies, and inconsistency due to the substantial variability of results.

The main findings of this review are that.

- **The overall group results are presented as a HSROC curve.** The diagnostic odds ratio (DOR) was 14.12 (95% confidence interval (CI) 11.17 to 17.84). In the absence of clinical consensus, we elected to report sensitivities at fixed values of specificity (median, upper quartile) as a means of expressing numerical quantities from the curve. This is in preference to using the average values of sensitivity and specificity which do not correspond to any particular threshold. The estimated points for sensitivity are 0.70 (95% CI 0.64 to 0.75) and 0.60 (95% CI 0.54 to 0.65), this is when applied at a fixed specificity of 0.78 and 0.90 (Summary of findings 1). There is a degree of non-independence of data in this analysis, as some studies provided multiple datasets. For a cohort of 1000 tooth sites or surfaces with a prevalence of enamel caries of 57% (the median prevalence observed in studies included in the meta-analysis), the sensitivity of 0.70 at a fixed specificity of 0.78 would result in 171 tooth sites not being identified as having early caries when caries was present (false negatives) and 95 tooth sites being identified as having caries when they did not (false positives). The consequences of these misclassifications are concerning, and all interventions have a cost at a patient or system level. A false positive classification for enamel caries would typically result in the application of topical fluoride or other minimally invasive treatments. A false negative classification implies that patients who require treatment would not receive it. Given the recall period for routine dental examinations and the slow-growing nature of the disease, the clinician may be reassured that the lesion could be identified at the patient's next appointment. The prevalence of enamel caries applied to this scenario is potentially inflated, due to many of the included studies being based on extracted teeth. In studies that employed an enhanced visual reference standard, based in either a school, primary care, or hospital setting, the median prevalence is lower at 47%.
- **There is no statistically significant difference in the accuracy of red, blue, or green fluorescence-based devices.** 84 (74%) of the available datasets allowed us to assess red fluorescence at the level of enamel caries, with 21 (18%) for blue, and 9 (8%) for green fluorescence, respectively. There was considerable heterogeneity of results within each of these subgroups that is reflective of the different reference standards, the prevalence of caries into dentine, tooth surface and dentition. A formal comparison of the fluorescence-based devices indicated that there was no difference in accuracy according to the category of the device (P = 0.14).
- **Studies with a higher proportion of observations with caries into dentine reported higher accuracy than studies with low and medium prevalence.** We considered the prevalence of caries into dentine to be important due to the potential for sensitivity and specificity to be inflated through the inclusion of large numbers of tooth surfaces with more advanced lesions obviously into dentine or frankly cavitated. These could be considerably more straightforward to detect, and therefore the inflation of accuracy estimates would occur. The investigation of the covariate of high prevalence (≥ 35%) versus medium (15% to 34%) and low (15%) prevalence concurred that this was occurring in the data gathered from the included studies, formal testing found that this difference was not statistically significant however (P = 0.32).
- **There is no meaningful difference in the accuracy of studies performed in vitro and in vivo.** The majority of studies were conducted on extracted teeth (in vitro) using a reference standard of histology, as opposed to teeth in situ conducted in

a clinical setting (in vivo). The results from the in vitro studies are essential for determining the validity of devices but do not truly inform us of the applicability of using these devices on patients in a general dental practice setting. Detecting disease in an in vitro setting can be assumed to be more straightforward than in a clinical setting as the challenges of accessing the tooth surfaces in the oral cavity, the complexity of soft tissues, or other teeth impeding the view, are largely eliminated. The evaluation of extracted teeth also facilitates the use of histology as a reliable and accurate reference standard; more recently microcomputed tomography (microCT) has also been used with some confidence as a reference standard, although this was not the case for any of the included studies. Since it is not feasible to extract and section healthy teeth and subject them to a histological reference standard, clinical studies have circumvented this issue by using enhanced visual examination or radiographs as effectively imperfect reference standards. The comparative accuracy of in vitro and in vivo study designs can be assessed by investigating the two most frequently used reference standards of histology (78 datasets, 68%) and enhanced visual assessment (25 datasets, 22%). Whilst the DOR was highest for enhanced visual examination as a reference standard formal comparison found no difference in accuracy (P = 0.06).

- **Diagnostic accuracy was higher for occlusal surfaces.** The majority of studies evaluated either occlusal (89 datasets) or proximal surfaces (18 sets). Some concern has been expressed that fluorescence-based devices are limited in their ability to detect proximal caries, as the excitation (laser) light needs to make direct contact with the tooth surface. If another tooth obstructs the excitation then the performance of the device will be suboptimal. There is no evidence that fluorescence-based devices show greater accuracy in detecting caries on occlusal surfaces than proximal surfaces (P = 0.62).
- **Studies on permanent teeth suggest greater accuracy over primary teeth when using fluorescence devices.** The distinction between the primary, mixed, and permanent dentition is of importance too. The detection of enamel caries may be of greater clinical importance in primary teeth as the depth of enamel is less than that of permanent teeth, and early caries could lead to more severe decay with greater expedience than would be witnessed in permanent teeth. However, the retention of permanent teeth throughout a person's lifetime is also important. Despite caries being seen as a slow-growing disease, the need for prevention in permanent teeth is also important. The results of the meta-analysis suggest that fluorescence devices may have greater accuracy in detecting caries in primary teeth, although this is not statistically significant (P = 0.19).
- **Devices that evaluated multiple sites on a tooth's surface showed a lower accuracy than those that evaluated a single site per tooth.** The assessment of multiple tooth sites introduces dependency, and a single underlying or hidden lesion could influence multiple sites. 24 of the 114 datasets in the meta-analysis reported multiple sites per tooth, however, five of these reported proximal surfaces where it would be less likely that this effect would occur. The results of the meta-analysis suggest that single point assessments may be more accurate, however, this was not statistically significant (P = 0.06). Two common methods were used when collecting a single site result per tooth, particularly when applied to the occlusal surface. Firstly where the device was passed over the tooth surface and the highest number from the device recorded, and secondly where the device was applied three times and the mean of the three recordings was taken.

Strengths and weaknesses of the review

The strengths of this Cochrane Review are the completion of a comprehensive literature search and rigorous application of methodology which ensured that all screening, inclusion decisions, and data extraction were performed in duplicate and with clinical input. Unlike many diagnostic test accuracy (DTA) systematic reviews, we did not restrict our inclusion criteria to studies presenting data in a 2 x 2 format, and this has enabled us to highlight the issue of incomplete reporting of outcome data and the inadequate reporting in primary DTA studies. We contacted study authors where necessary to ensure that we could obtain data for as many studies as possible. Further, we used a clear and reproducible process for methodological decision making.

The substantial number of included studies facilitated meta-analysis. The primary analysis was conducted using hierarchical summary receiver operating characteristic (HSROC) curves rather than the Moses-Littenberg method which has been used in other caries DTA reviews, and which has been shown to perform poorly in comparison to hierarchical approaches (Dinnes 2016). An HSROC approach was undertaken as opposed to the bivariate method due to the variation of thresholds employed between sound and carious tooth surfaces in the included studies. The quoted sensitivities and specificities used to calculate the natural frequencies should therefore be interpreted cautiously.

This review comprises a substantial number of primary studies. Bader 2004 completed a review of fluorescence devices, and this Cochrane Review is a significant update that broadens the remit of the earlier review to include visual and radiographs reference standards in addition to histology. This DTA systematic review has substantially increased the number of included studies from 25 (Bader 2004) and 73 in a more recent review (Gimenez 2013) to 133 (79 studies included in the meta-analysis) in this review. The use of HSROC methodology is an important component of this DTA systematic review. Gimenez 2013 did not use the hierarchical model, although our conclusion is similar - that fluorescence devices show improved results in more severe caries, but that the accuracy of devices is similar across different tooth surfaces. Our review also focuses on the target condition of early enamel lesions which has the potential to inform clinicians on the decision to intervene earlier in the disease process with preventive or minimally invasive treatments rather than operative.

The main weakness of the review is the substantial volume of studies with incomplete outcome data. 55 of the 133 included studies provided insufficient information to enable us to construct or compute a 2 x 2 table. Many studies did not present the numbers of true positives, true negatives, false positives, and false negatives at the enamel threshold. Rather, they reported sensitivity, specificity, and area under the curve as their primary results. This did not allow us to include a study in the meta-analysis unless the prevalence of caries at the enamel threshold was reported, enabling the construction of the required 2 x 2 table of outcomes.

A significant source of bias in many of the studies was that the participants or teeth were selected, with the risk that teeth were selected teeth that made caries detection more straightforward, with resulting inflation of sensitivity and/or specificity values. When planning the meta-analysis, it became apparent that an argument could be created to subgroup by in vitro and in vivo studies, or by index test, or by the reference standard. We decided to allow the primary meta-analysis to remain as a single complete dataset and to investigate the effects of these factors through meta-regression, and to allow the results of this analysis to guide the remainder of the meta-analysis.

The inclusion criteria were selected to ensure that the focus of the review was on the detection of early caries or caries limited to enamel. However, with the best of intentions studies could easily attempt to recruit sound or non-cavitated teeth but when investigated with a thorough/complete reference standard it became apparent that when viewed during participant selection, surfaces harboured dentinal caries. The concern of the review team was that if studies intentionally recruited dentinal lesions, then there would be a simplification of the detection and diagnostic decision as a lesion which was validated and reached dentine is generally easier to observe than an early lesion which is limited to the enamel. A further complication arose where some studies were poorly reported or lacked clarity on the selection criteria that they imposed on their sample. We took the position that unless the authors clearly stated that frank or dentinal cavities were intentionally included, then we were unable to exclude the study from the review. The result of this decision has been difficult to apply consistently, and consequently, we may have excluded some well-reported studies due to their clarity of reporting, whereas studies which intentionally included dentinal lesions, but failed to report this inclusion, were included. We accept this may leave the review open to some criticism, and we would reiterate that this review intended to synthesise the evidence on early lesions. The inclusion of more advanced lesions that are obviously into dentine or frankly cavitated does not fit the remit of this review. Analysis of the prevalence of caries at the dentinal level enabled us to investigate this assertion which results confirmed.

Some studies purposefully investigated the most accurate threshold, using the study data and ROC curve to determine the optimum threshold to maximise values of sensitivity or specificity or both. The focus of our review, however, was on the accuracy of these devices when used by general dentists, which requires the use of a pre-defined threshold. The reporting of results according to optimised data-driven thresholds is problematic as the observed sensitivity and specificity values will be higher in these studies than those applying pre-determined thresholds, the thresholds selected by these studies may not be generalisable to other patient populations. Although useful, such studies may have limited relevance to our research question. Another area of concern arose when the reference standard was histology and studies did not report whether the same examiners conducted the index test and reference standard assessments. This issue was logged in the characteristics of included studies tables, but our interpretation was that this would not affect the judgement of the reference standard as it was hard to see how an examiner would remember the results of the fluorescence devices and recall it during the examination of a sectioned tooth. A final area of concern was the effect of the chosen threshold between the sound, enamel caries, and dentinal caries. For example, the thresholds used for the DIAGNOdent device to differentiate between sound and dentinal caries ranged from 2 to 20 so the results of one study could be reassessed according to other thresholds and very different results obtained. As the HSROC approach models threshold effects no further assessment was required.

Applicability of findings to the review question

There are concerns regarding the clinical applicability of the findings of this review resulting from the fact that 68% of the datasets are based on in vitro studies and therefore not conducted in a setting which is representative of the general dental setting. Until a more perfect reference standard for safe use in vivo is developed, this is likely to be the status quo. Developments in the use of 3D technology in vitro (microCT) and in vivo (cone-beam CT) may go some way to improve upon these concerns.

AUTHORS' CONCLUSIONS

Implications for practice

We intended that the results of this review be directly applicable to the general dental practitioner. Ideally, clinicians would have all diagnostic test or devices available to them and use the most appropriate according to the clinical scenario. This is not possible for most dental practices who have finite resources and existing infrastructure which would almost always feature a radiographic device to support the conventional oral examination. The question remains to clinicians whether the utilisation of a fluorescence device provides sufficient benefits to justify the cost. There is considerable variation in the performance of the fluorescence-based devices included in this Cochrane Review that could not be explained by the different wavelengths of the devices assessed, or by participant or study characteristics. Blue and green fluorescence-based devices appeared to outperform red fluorescence-based devices, but this difference was not supported by the results of a formal statistical comparison. There are concerns that these results may be confounded by stain, and that the lower number of studies included for some blue fluorescence devices means that further research into the accuracy of these devices may be warranted. The reproducibility of the devices was beyond the scope of this review, but one important, clinically useful application could be the use of these devices over multiple time points to monitor lesions or even to quantify lesion severity to justify any intervention. Clinicians will always perform a visual examination but may well look to an adjunct to provide validation or confirmation of their decision. Due to the low certainty of the evidence from studies included in this review, considerable uncertainty remains regarding the accuracy of fluorescence-based devices for early caries detection.

Bader 2004 recommended that fluorescent devices should not be used in isolation and based on the certainty of the evidence there is little to challenge this recommendation. Despite the reasonably high sensitivity and specificity estimates, we cannot envisage a scenario where a clinician would carry out a clinical examination without performing a thorough visual diagnosis, and with development future fluorescence-based devices may support the clinician in confirming the status of uncertain or difficult to diagnose teeth.

Implications for research

As is highlighted by the number of studies included in this review which did not report data in a useable format, it is of vital importance that future research studies report the data in a clear concise method and following the STARD checklist (Bossuyt 2003; Bossuyt 2015), ideally with a cross-tabulation of the index test and reference standard with a minimum requirement of three categories of each which could be classified as sound/caries free, early/enamel caries, advanced/dentine caries. Many studies subdivided these latter two categories into inner and outer enamel/dentine caries, and this allowed us to extract true-positive, false-positive, false-negative, and true-positive results.

Importantly, future studies should be aware of the importance of sampling participants using consecutive or random sampling. This should serve to minimise the bias which originates from the selection of teeth in which early caries is either easier or more difficult to detect. Sensitivity analysis suggested that sensitivity and specificity could be overestimated by failing to use random or consecutive sampling. Studies should also specify the test positivity thresholds a priori rather than selecting the threshold which maximises estimates of sensitivity and specificity, ideally using manufacturer recommended thresholds or those validated in previous research studies. Studies may be conducted to determine the most accurate thresholds for a given population. We would recommend that studies such as these report the manufacturer recommended thresholds in addition to the maximised thresholds to facilitate a comparison between the two and allow for analysis in future reviews.

When designing the ideal study for future research, it is important to consider the single study that we judged to be at low risk of bias and low concern across all domains for Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2). This study identified children that required a tooth extraction, which enabled the index test to be conducted in the clinical setting, and a histology reference standard once the tooth had been extracted. Future studies could look at the potential of fluorescence devices to be used in combination with other technologies and to make direct comparisons between their use at different points of the disease spectrum, i.e. general practice: seemingly asymptomatic, low/high need, irregular attenders, previously diseased participants. Given the potential utility of the devices in aiding the clinician to confirm borderline cases where the clinician is uncertain of the true disease state, a study could be designed which investigates only those sites which have a degree of uncertainty.

Randomised controlled trials could be beneficial by investigating the longer-term effects of using the fluorescence devices for detection, diagnosis, and monitoring to identify whether they aid the prevention of disease through active preventative interventions.

ACKNOWLEDGEMENTS

This series of Cochrane Reviews was funded by the UK National Institute for Health Research (NIHR) Cochrane Programme Grant Scheme (Project: 16/114/23). We would like to thank Anne Littlewood (Information Specialist, Cochrane Oral Health) for her advice on the search strategy and conducting the search of the literature, and Luisa M Fernandez Mauleffinch (Managing Editor and Copy Editor, Cochrane Oral Health) for her assistance in facilitating this review. We thank Richard Hogan and Iain Pretty for their advice during protocol development and the interpretation of findings; Associate Professor KR Ekstrand, and J Bader (Emeritus Professor, UNC School of Dentistry, Chapel Hill North Carolina, USA) for their feedback on the protocol; Jennifer Hilgart, J Bader, Alonso Carrasco-Labra (Senior Director, Department of Evidence Synthesis and Translation Research, ADA Science & Research Institute, LLC), and the Cochrane Diagnostic Test Accuracy Editorial Team for their feedback on the review. Also Alex Sutton and Suzanne Freeman from the NIHR Complex Review Support Unit for their support on this review. For translation of primary studies we wish to thank Margriet van Baar, Fang Hua, Szimonetta Lohner, Anette Bluemle, Karin Bischoff, and Maddalena Manfredi.

REFERENCES

References to studies included in this review

Achilleos 2013 *{published data only}*

Achilleos EE, Rahiotis C, Kakaboura A, Vougiouklakis G. Evaluation of a new fluorescence-based device in the detection of incipient occlusal caries lesions. *Lasers in Medical Science* 2013;**28**:193-201.

Akarsu 2006 *{published data only}*

Akarsu S, Koprulu H. In vivo comparison of the efficacy of DIAGNOdent by visual inspection and radiographic diagnostic techniques in the diagnosis of occlusal caries. *Journal of Clinical Dentistry* 2006;**17**(3):53-8.

Aktan 2012 *{published data only}*

Aktan AM, Cebe MA, Ciftci ME, Sirin Karaarslan E. A novel LED-based device for occlusal caries detection. *Lasers in Medical Science* 2012;**27**:1157-63.

Almosa 2014 *{published data only}*

Almosa NA, Lundgren T, Aldrees AM, Birkhed D, Kjellberg H. Diagnosing the severity of buccal caries lesions in governmental and private orthodontic patients at debonding, using the ICDAS-II and the DIAGNOdent pen. *Angle Orthodontist* 2014;**84**:430-6.

Alomari 2015 *{published data only}*

Alomari QD, Qudeimat MA, Khalaf ME, Al-Tarakemah Y. The effect of combining radiographs and DIAGNOdent with visual examination on detection and treatment decisions of noncavitated occluso-dentinal caries. *Operative Dentistry* 2015;**40**:313-21.

Alwas-Danowska 2002 *{published data only}*

Alwas-Danowska HM, Plasschaert AJ, Suliborski S, Verdonschot EH. Reliability and validity issues of laser fluorescence measurements in occlusal caries diagnosis. *Journal of Dentistry* 2002;**30**:129-34.

Angnes 2005 *{published data only}*

Angnes V, Angnes G, Batisttella M, Grande RH, Loguercio AD, Reis A. Clinical effectiveness of laser fluorescence, visual inspection and radiography in the detection of occlusal caries. *Caries Research* 2005;**39**:490-5.

Anttonen 2003 *{published data only}*

Anttonen V, Seppa L, Hausen H. Clinical study of the use of the laser fluorescence device DIAGNOdent for detection of occlusal caries in children. *Caries Research* 2003;**37**:17-23.

Apostolopoulou 2009 *{published data only}*

Apostolopoulou D, Lagouvardos P, Kavvadia K, Papagiannoulis L. Histological validation of a laser fluorescence device for occlusal caries detection in primary molars. *European Archives of Paediatric Dentistry* 2009;**10 Suppl 1**:11-5.

Arslan 2014 *{published data only}*

Arslan U, Karaagaoglu E, Ozkan G, Kanli A. Evaluation of diagnostic tests using information theory for multi-class diagnostic problems and its application for the detection of occlusal caries lesions. *Balkan Medical Journal* 2014;**31**(3):214-8.

Attrill 2001 *{published data only}*

Attrill DC, Ashley PF. Occlusal caries detection in primary teeth: a comparison of DIAGNOdent with conventional methods. *British Dental Journal* 2001;**190**:440-3.

Bahrololoomi 2015 *{published data only}*

Bahrololoomi Z, Ezoddini F, Halvani N. Comparison of radiography, laser fluorescence and visual examination for diagnosing incipient occlusal caries of permanent first molars. *Journal of Dentistry (Tehran, Iran)* 2015;**12**(5):324-32.

Bamzahim 2002 *{published data only}*

Bamzahim M, Shi XQ, Angmar-Mansson B. Occlusal caries detection and quantification by DIAGNOdent and Electronic Caries Monitor: in vitro comparison. *Acta Odontologica Scandinavica* 2002;**60**:360-4.

Bamzahim 2004 *{published data only}*

Bamzahim M, Shi XQ, Angmar-Månsson B. Secondary caries detection by DIAGNOdent and radiography: a comparative in vitro study. *Acta Odontologica Scandinavica* 2004;**62**(1):61-4.

Barberia 2008 *{published data only}*

Barberia E, Maroto M, Arenas M, Silva CC. A clinical study of caries diagnosis with a laser fluorescence system. *Journal of the American Dental Association* 2008;**139**:572-9.

Baseren 2003 *{published data only}*

Baseren NM, Gokalp S. Validity of a laser fluorescence system (DIAGNOdent) for detection of occlusal caries in third molars: an in vitro study. *Journal of Oral Rehabilitation* 2003;**30**:1190-4.

Bengtson 2005 *{published data only}*

Bengtson AL, Gomes AC, Mendes FM, Cichello LR, Bengtson NG, Pinheiro SL. Influence of examiner's clinical experience in detecting occlusal caries lesions in primary teeth. *Pediatric Dentistry* 2005;**27**:238-43.

Bittar 2012 *{published data only}*

Bittar DG, Gimenez T, Morais CC, De Benedetto MS, Braga MM, Mendes FM. Influence of moisture and plaque on the performance of a laser fluorescence device in detecting caries lesions in primary teeth. *Lasers in Medical Science* 2012;**27**:1169-74.

Bizhang 2016 *{published data only}*

Bizhang M, Wollenweber N, Singh-Husgen P, Danesh G, Zimmer S. Pen-type laser fluorescence device versus bitewing radiographs for caries detection on approximal surfaces. *Head & Face Medicine* 2016;**12**:30.

Boston 2003 *{published data only}*

Boston DW. Initial in vitro evaluation of DIAGNOdent for detecting secondary carious lesions associated with

resin composite restorations. *Quintessence International* 2003;**34**:109-16.

Bozdemir 2013 *{published data only}*

Bozdemir E, Karaarslan ES, Ozsevik AS, Ata Cebe M, Aktan AM. In vivo performance of two devices for occlusal caries detection. *Photomedicine and Laser Surgery* 2013;**31**:322-7.

Braga 2006 *{published data only}*

Braga MM, Mendes FM, Martins CR, Imparato JC. Effect of the calibration method of a laser fluorescence device for detecting occlusal caries in primary molars. *Pediatric Dentistry* 2006;**28**:451-4.

Braga 2007 *{published data only}*

Braga MM, Mendes FM, Imparato JC, Rodrigues CR. Effect of cut-off points on performance of laser fluorescence for detecting occlusal caries. *Journal of Clinical Pediatric Dentistry* 2007;**32**:33-6.

Braga 2008 *{published data only}*

Braga M, Nicolau J, Rodrigues CR, Imparato JC, Mendes FM. Laser fluorescence device does not perform well in detection of early caries lesions in primary teeth: an in vitro study. *Oral Health & Preventive Dentistry* 2008;**6**:165-9.

Braga 2009 *{published data only}*

Braga MM, Morais CC, Nakama RCS, Leamari VM, Siqueira WL, Mendes FM. In vitro performance of methods of approximal caries detection in primary molars. *Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology and Endodontology* 2009;**108**:e35-41.

Burin 2005 *{published data only}*

Burin C, Burin C, Loguercio AD, Grande RH, Reis A. Occlusal caries detection: a comparison of a laser fluorescence system and conventional methods. *Pediatric Dentistry* 2005;**27**:307-12.

Bussaneli 2015 *{published data only}*

Bussaneli DG, Restrepo M, Boldieri T, Pretel H, Mancini MW, Santos-Pinto L, et al. Assessment of a new infrared laser transillumination technology (808 nm) for the detection of occlusal caries - an in vitro study. *Lasers in Medical Science* 2015;**30**(7):1873-9.

Bussaneli 2015a *{published data only}*

Bussaneli DG, Restrepo M, Boldieri T, Albertoni TH, Santos-Pinto L, Cordeiro RCL. Proximal caries lesion detection in primary teeth: does this justify the association of diagnostic methods? *Lasers in Medical Science* 2015;**30**(9):2239-44.

Castilho 2016 *{published data only}*

Castilho LS, Cotta FV, Bueno AC, Moreira AN, Ferreira EF, Magalhaes CS. Validation of DIAGNOdent laser fluorescence and the International Caries Detection and Assessment System (ICDAS) in diagnosis of occlusal caries in permanent teeth: an in vivo study. *European Journal of Oral Sciences* 2016;**124**:188-94.

Chawla 2012 *{published data only}*

Chawla N, Messer LB, Adams GG, Manton DJ. An in vitro comparison of detection methods for approximal carious lesions in primary molars. *Caries Research* 2012;**46**(2):161-9.

Chen 2012 *{published data only}*

Chen J, Qin M, Ma W, Ge L. A clinical study of a laser fluorescence device for the detection of approximal caries in primary molars. *International Journal of Paediatric Dentistry* 2012;**22**:132-8.

Chong 2003 *{published data only}*

Chong MJ, Seow WK, Purdie DM, Cheng E, Wan V. Visual-tactile examination compared with conventional radiography, digital radiography, and Diagnodent in the diagnosis of occlusal occult caries in extracted premolars. *Pediatric Dentistry* 2003;**25**:341-9.

Cinar 2013 *{published data only}*

Cinar C, Atabek D, Odabas ME, Olmez A. Comparison of laser fluorescence devices for detection of caries in primary teeth. *International Dental Journal* 2013;**63**:97-102.

Costa 2002 *{published data only}*

Costa AM, Yamaguti PM, De Paula LM, Bezerra AC. In vitro study of laser diode 655 nm diagnosis of occlusal caries. *Journal of Dentistry for Children* 2002;**69**(3):249-53, 233.

Costa 2007 *{published data only}*

Costa AM, Bezzerra AC, Fuks AB. Assessment of the accuracy of visual examination, bite-wing radiographs and DIAGNOdent on the diagnosis of occlusal caries. *European Archives of Paediatric Dentistry* 2007;**8**:118-22.

Diniz 2009 *{published data only}*

Diniz MB, Rodrigues JA, de Paula AB, Cordeiro R de C. In vivo evaluation of laser fluorescence performance using different cut-off limits for occlusal caries detection. *Lasers in Medical Science* 2009;**24**:295-300.

Diniz 2011 *{published data only}*

Diniz MB, Sciasci P, Rodrigues JA, Lussi A, Cordeiro RCL. Influence of different professional prophylactic methods on fluorescence measurements for detection of occlusal caries. *Caries Research* 2011;**45**(3):264-8.

Diniz 2012 *{published data only}*

Diniz MB, Boldieri T, Rodrigues JA, Santos-Pinto L, Lussi A, Cordeiro RC. The performance of conventional and fluorescence-based methods for occlusal caries detection: an in vivo study with histologic validation. *Journal of the American Dental Association* 2012;**143**(4):339-50.

Diniz 2019 *{published data only}*

Diniz MB, Campos PH, Wilde S, Cordeiro RCL, Zandona AGF. Performance of light-emitting diode device in detecting occlusal caries in the primary molars. *Lasers in Medical Science* 2019;**34**(9):1235-41.

Duruturk 2011 *{published data only}*

Duruturk L, Ciftci A, Baharoglu S, Oztuna D. Clinical evaluation of DIAGNOdent in detection of occlusal caries in newly erupted

noncavitated first permanent molars in caries-active children. *Operative Dentistry* 2011;**36**:348-55.

El-Housseiny 2001 *{published data only}*

El-Housseiny AA, Jamjoum H. Evaluation of visual, explorer, and a laser device for detection of early occlusal caries. *Journal of Clinical Pediatric Dentistry* 2001;**26**:41-8.

Feng 2005 *{published data only}*

Feng Y, Yin W, Zhang YY, Zhang B, Hu DY. Comparison of primary caries detection on smooth surface in the maxillary anterior teeth using QLF, digital photo and visual diagnosis. *Shanghai Kou Qiang Yi Xue [Shanghai Journal of Stomatology]* 2005;**14**(6):565-8.

Ferreira 1998 *{published data only}*

Ferreira Zandona AG, Analoui M, Beiswanger BB, Isaacs RL, Kafrawy AH, Eckert GJ, et al. An in vitro comparison between laser fluorescence and visual examination for detection of demineralization in occlusal pits and fissures. *Caries Research* 1998;**32**:210-8.

Ferreira 2008 *{published data only}*

Ferreira JM, Silva MF, Oliveira AF, Sampaio FC. Evaluation of different methods for monitoring incipient carious lesions in smooth surfaces under fluoride varnish therapy. *International Journal of Paediatric Dentistry* 2008;**18**:300-5.

Francescut 2003 *{published data only}*

Francescut P, Lussi A. Correlation between fissure discoloration, Diagnodent measurements, and caries depth: an in vitro study. *Pediatric Dentistry* 2003;**25**:559-64.

Fung 2004 *{published data only}*

Fung L, Smales R, Ngo H, Mount G. Diagnostic comparison of three groups of examiners using visual and laser fluorescence methods to detect occlusal caries in vitro. *Australian Dental Journal* 2004;**49**:67-71.

Ghaname 2010 *{published data only}*

Ghaname ES, Ritter AV, Heymann HO, Vann WF Jr, Shugars DA, Bader JD. Correlation between laser fluorescence readings and volume of tooth preparation in incipient occlusal caries in vitro. *Journal of Esthetic & Restorative Dentistry* 2010;**22**:31-9.

Goel 2009 *{published data only}*

Goel A, Chawla HS, Gauba K, Goyal A. Comparison of validity of DIAGNOdent with conventional methods for detection of occlusal caries in primary molars using the histological gold standard: an in vivo study. *Journal of the Indian Society of Pedodontics and Preventive Dentistry* 2009;**27**(4):227-34.

Graye 2012 *{published data only}*

Graye M, Markowitz K, Strickland M, Guzy G, Burke M, Houpt M. In vitro evaluation of the Spectra early caries detection system. *Journal of Clinical Dentistry* 2012;**23**:1-6.

Heinrich-Weltzien 2003 *{published data only}*

Heinrich-Weltzien R, Kuhnisch J, Oehme T, Ziehe A, Stosser L, Garcia-Godoy F. Comparison of different DIAGNOdent cut-off limits for in vivo detection of occlusal caries. *Operative Dentistry* 2003;**28**(6):672-80.

Hibst 2001 *{published data only}*

Hibst R, Paulus R, Lussi A. Detection of occlusal caries by laser fluorescence: basic and clinical investigations. *Medical Laser Application* 2001;**16**(3):205-13.

Huth 2008 *{published data only}*

Huth KC, Neuhaus KW, Gygax M, Bucher K, Crispin A, Paschos E, et al. Clinical performance of a new laser fluorescence device for detection of occlusal caries lesions in permanent molars. *Journal of Dentistry* 2008;**36**(12):1033-40.

Huth 2010 *{published data only}*

Huth KC, Lussi A, Gygax M, Thum M, Crispin A, Paschos E, et al. In vivo performance of a laser fluorescence device for the approximal detection of caries in permanent molars. *Journal of Dentistry* 2010;**38**(12):1019-26.

Iranzo-Cortes 2017 *{published data only}*

Iranzo-Cortes JE, Terzic S, Montiel-Company JM, Almerich-Silla JM. Diagnostic validity of ICDAS and DIAGNOdent combined: an in vitro study in pre-cavitated lesions. *Lasers in Medical Science* 2017;**32**(3):543-8.

Jablonski-Momeni 2011 *{published data only}*

Jablonski-Momeni A, Schipper HM, Rosen SM, Heinzel-Gutenbrunner M, Roggendorf MJ, Stoll R, et al. Performance of a fluorescence camera for detection of occlusal caries in vitro. *Odontology* 2011;**99**(1):55-61.

Jablonski-Momeni 2012 *{published data only}*

Jablonski-Momeni A, Rosen SM, Schipper HM, Stoll R, Roggendorf MJ, Heinzel-Gutenbrunner M, et al. Impact of measuring multiple or single occlusal lesions on estimates of diagnostic accuracy using fluorescence methods. *Lasers in Medical Science* 2012;**27**(2):343-52.

Jablonski-Momeni 2012a *{published data only}*

Jablonski-Momeni A, Stucke J, Steinberg T, Heinzel-Gutenbrunner M. Use of ICDAS-II, fluorescence-based methods, and radiography in detection and treatment decision of occlusal caries lesions: an in vitro study. *International Journal of Dentistry* 2012;**2012**:371595.

Jablonski-Momeni 2014 *{published data only}*

Jablonski-Momeni A, Heinzel-Gutenbrunner M, Klein SM. In vivo performance of the VistaProof fluorescence-based camera for detection of occlusal lesions. *Clinical Oral Investigations* 2014;**18**(7):1757-62.

Jablonski-Momeni 2016 *{published data only}*

Jablonski-Momeni A, Heinzel-Gutenbrunner M, Vill G. Use of a fluorescence-based camera for monitoring occlusal surfaces of primary and permanent teeth. *International Journal of Paediatric Dentistry* 2016;**26**(6):448-56.

Jeon 2004 *{published data only}*

Jeon RJ, Han C, Mandelis A, Sanchez V, Abrams SH. Diagnosis of pit and fissure caries using frequency-domain infrared

photothermal radiometry and modulated laser luminescence. *Caries Research* 2004;**38**(6):497-513.

Jung 2018 *{published data only}*

Jung EH, Lee ES, Jung HI, Kang SM, de Josselin de Jong E, Kim BI. Development of a fluorescence-image scoring system for assessing noncavitated occlusal caries. *Photodiagnosis and Photodynamic Therapy* 2018;**21**:36-42.

Kavvadia 2008 *{published data only}*

Kavvadia K, Lagouvardos P. Clinical performance of a diode laser fluorescence device for the detection of occlusal caries in primary teeth. *International Journal of Paediatric Dentistry* 2008;**18**(3):197-204.

Kavvadia 2012 *{published data only}*

Kavvadia K, Lagouvardos P, Apostolopoulou D. Combined validity of DIAGNOdentTM and visual examination for in vitro detection of occlusal caries in primary molars. *Lasers in Medical Science* 2012;**27**(2):313-9.

Kesler 2003 *{published data only}*

Kesler G, Masychev V, Sokolovsky A, Alexandrov M, Kesler A, Koren R. Photon undulatory non-linear conversion diagnostic method for caries detection: a pilot study. *Journal of Clinical Laser Medicine & Surgery* 2003;**21**(4):209-17.

Kim 2017 *{published data only}*

Kim ES, Lee ES, Kang SM, Jung EH, de Josselin de Jong E, Jung HI, et al. A new screening method to detect proximal dental caries using fluorescence imaging. *Photodiagnosis and Photodynamic Therapy* 2017;**20**:257-62.

Ko 2015 *{published data only}*

Ko HY, Kang SM, Kim HE, Kwon HK, Kim BI. Validation of quantitative light-induced fluorescence-digital (QLF-D) for the detection of approximal caries in vitro. *Journal of Dentistry* 2015;**43**(5):568-75.

Kockanat 2017 *{published data only}*

Kockanat A, Unal M. In vivo and in vitro comparison of ICDAS II, DIAGNOdent pen, CarieScan PRO and SoproLife camera for occlusal caries detection in primary molar teeth. *European Journal of Paediatric Dentistry* 2017;**18**(2):99-104.

Kouchaji 2012 *{published data only}*

Kouchaji C. Comparison between a laser fluorescence device and visual examination in the detection of occlusal caries in children. *Saudi Dental Journal* 2012;**24**(3-4):169-74.

Krause 2007 *{published data only}*

Krause F, Jepsen S, Braun A. Comparison of two laser fluorescence devices for the detection of occlusal caries in vivo. *European Journal of Oral Sciences* 2007;**115**(4):252-6.

Kucukyilmaz 2015 *{published data only}*

Kucukyilmaz E, Sener Y, Botsali MS. In vivo and in vitro performance of conventional methods, DIAGNOdent, and an electronic caries monitor for occlusal caries detection in primary teeth. *Pediatric Dentistry* 2015;**37**(4):E14-22.

Kuhnisch 2006 *{published data only}*

Kuhnisch J, Ifland S, Tranaeus S, Angmar-Mansson B, Hickel R, Stosser L, et al. Establishing quantitative light-induced fluorescence cut-offs for the detection of occlusal dentine lesions. *European Journal of Oral Sciences* 2006;**114**(6):483-8.

Kuhnisch 2007 *{published data only}*

Kuhnisch J, Ifland S, Tranaeus S, Hickel R, Stosser L, Heinrich-Weltzien R. In vivo detection of non-cavitated caries lesions on occlusal surfaces by visual inspection and quantitative light-induced fluorescence. *Acta Odontologica Scandinavica* 2007;**65**(3):183-8.

Kuhnisch 2008 *{published data only}*

Kuhnisch J, Berger S, Goddon I, Senkel H, Pitts N, Heinrich-Weltzien R. Occlusal caries detection in permanent molars according to WHO basic methods, ICDAS II and laser fluorescence measurements. *Community Dentistry and Oral Epidemiology* 2008;**36**(6):475-84.

Lee 2018 *{published data only}*

Lee HS, Kim SK, Park SW, Jong E de Josselin de, Kwon HK, Jeong SH, et al. Caries detection and quantification around stained pits and fissures in occlusal tooth surfaces with fluorescence. *Journal of Biomedical Optics* 2018;**23**(9):1-7.

Li 2006 *{published data only}*

Li X, Fan X, Jia SH, Hu DY. Clinical study of use of laser fluorescence for detecting occlusal caries in deciduous teeth. *Hua Xi Kou Qiang Yi Xue Za Zhi [West China Journal of Stomatology]* 2006;**24**(1):36-8.

Lussi 1999 *{published data only}*

Lussi A, Imwinkelried S, Pitts N, Longbottom C, Reich E. Performance and reproducibility of a laser fluorescence system for detection of occlusal caries in vitro. *Caries Research* 1999;**33**(4):261-6.

Lussi 2001 *{published data only}*

Lussi A, Megert B, Longbottom C, Reich E, Francescut P. Clinical performance of a laser fluorescence device for detection of occlusal caries lesions. *European Journal of Oral Sciences* 2001;**109**(1):14-9.

Lussi 2003 *{published data only}*

Lussi A, Francescut P. Performance of conventional and new methods for the detection of occlusal caries in deciduous teeth. *Caries Research* 2003;**37**(1):2-7.

Lussi 2005 *{published data only}*

Lussi A, Longbottom C, Gygax M, Braig F. Influence of professional cleaning and drying of occlusal surfaces on laser fluorescence in vivo. *Caries Research* 2005;**39**(4):284-6.

Lussi 2006 *{published data only}*

Lussi A, Hack A, Hug I, Heckenberger H, Megert B, Stich H. Detection of approximal caries with a new laser fluorescence device. *Caries Research* 2006;**40**(2):97-103.

Lussi 2006a *{published data only}*

Lussi A, Hellwig E. Performance of a new laser fluorescence device for the detection of occlusal caries in vitro. *Journal of Dentistry* 2006;**34**(7):467-71.

Mansour 2016 *{published data only}*

Mansour S, Ajdaharian J, Nabelsi T, Chan G, Wilder-Smith P. Comparison of caries diagnostic modalities: a clinical study in 40 subjects. *Lasers in Surgery and Medicine* 2016;**48**(10):924-8.

Manton 2007 *{published data only}*

Manton DJ, Messer LB. The effect of pit and fissure sealants on the detection of occlusal caries in vitro. *European Archives of Paediatric Dentistry* 2007;**8**(1):43-8.

Markowitz 2013 *{published data only}*

Markowitz K, Rosenfeld D, Peikes D, Guzy G, Rosivack G. Effect of pit and fissure sealants on caries detection by a fluorescent camera system. *Journal of Dentistry* 2013;**41**(7):590-9.

Markowitz 2015 *{published data only}*

Markowitz K, Gutta A, Merdad H E, Guzy G, Rosivack G. In vitro study of the diagnostic performance of the Spectra Caries Detection Aid. *Journal of Clinical Dentistry* 2015;**26**(1):17-22.

Matos 2011 *{published data only}*

Matos R, Novaes TF, Braga MM, Siqueira WL, Duarte DA, Mendes FM. Clinical performance of two fluorescence-based methods in detecting occlusal caries lesions in primary teeth. *Caries Research* 2011;**45**(3):294-302.

Mendes 2005 *{published data only}*

Mendes FM, Siqueira WL, Mazzitelli JF, Pinheiro SL, Bengtson AL. Performance of DIAGNOdent for detection and quantification of smooth-surface caries in primary teeth. *Journal of Dentistry* 2005;**33**(1):79-84.

Mendes 2006 *{published data only}*

Mendes FM, Ganzerla E, Nunes AF, Puig AV, Imparato JC. Use of high-powered magnification to detect occlusal caries in primary teeth. *American Journal of Dentistry* 2006;**19**(1):19-22.

Mendes 2012 *{published data only}*

Mendes FM, Novaes TF, Matos R, Bittar DG, Piovesan C, Gimenez T, et al. Radiographic and laser fluorescence methods have no benefits for detecting caries in primary teeth. *Caries Research* 2012;**46**(6):536-43.

Mepparambath 2014 *{published data only}*

Mepparambath R, S Bhat S, K Hegde S, Anjana G, Sunil M, Mathew S. Comparison of proximal caries detection in primary teeth between laser fluorescence and bitewing radiography: an in vivo study. *Jaypees International Journal of Clinical Pediatric Dentistry* 2014;**7**(3):163-7.

Mortensen 2018 *{published data only}*

Mortensen D, Hessing-Olsen I, Ekstrand KR, Twetman S. In-vivo performance of impedance spectroscopy, laser fluorescence, and bitewing radiographs for occlusal caries detection. *Quintessence International* 2018;**49**(4):293-9.

Muller-Bolla 2017 *{published data only}*

Muller-Bolla M, Joseph C, Pisapia M, Tramini P, Velly AM, Tassery H. Performance of a recent light fluorescence device for detection of occlusal carious lesions in children and adolescents. *European Archives of Paediatric Dentistry* 2017;**18**(3):187-95.

Neuhaus 2011 *{published data only}*

Neuhaus KW, Rodrigues JA, Hug I, Stich H, Lussi A. Performance of laser fluorescence devices, visual and radiographic examination for the detection of occlusal caries in primary molars. *Clinical Oral Investigations* 2011;**15**(5):635-41.

Novaes 2009 *{published data only}*

Novaes TF, Matos R, Braga MM, Imparato JC, Raggio DP, Mendes FM. Performance of a pen-type laser fluorescence device and conventional methods in detecting approximal caries lesions in primary teeth - in vivo study. *Caries Research* 2009;**43**(1):36-42.

Novaes 2010 *{published data only}*

Novaes TF, Matos R, Raggio DP, Imparato JC, Braga MM, Mendes FM. Influence of the discomfort reported by children on the performance of approximal caries detection methods. *Caries Research* 2010;**44**(5):465-71.

Novaes 2012 *{published data only}*

Novaes TF, Matos R, Celiberti P, Braga MM, Mendes FM. The influence of interdental spacing on the detection of proximal caries lesions in primary teeth. *Brazilian Oral Research* 2012;**26**(4):293-9.

Novaes 2012a *{published data only}*

Novaes TF, Matos R, Gimenez T, Braga MM, DE Benedetto MS, Mendes FM. Performance of fluorescence-based and conventional methods of occlusal caries detection in primary molars - an in vitro study. *International Journal of Paediatric Dentistry* 2012;**22**(6):459-66.

Novaes 2016 *{published data only}*

Novaes TF, Moriyama CM, De Benedetto MS, Kohara EK, Braga MM, Mendes FM. Performance of fluorescence-based methods for detecting and quantifying smooth-surface caries lesions in primary teeth: an in vitro study. *International Journal of Paediatric Dentistry* 2016;**26**(1):13-9.

Ouellet 2002 *{published data only}*

Ouellet A, Hondrum SO, Pietz DM. Detection of occlusal carious lesions. *General Dentistry* 2002;**50**(4):346-50.

Ozsevik 2015 *{published data only}*

Ozsevik AS, Kararslan ES, Aktan AM, Bozdemir E, Cebe F, Sari F. Effect of different contact materials on approximal caries detection by laser fluorescence and light-emitting diode devices. *Photomedicine and Laser Surgery* 2015;**33**(10):492-7.

Ozturk 2015 *{published data only}*

Ozturk E, Sinanoglu A. Histological validation of cone-beam computed tomography versus laser fluorescence and conventional diagnostic methods for occlusal caries detection. *Photomedicine and Laser Surgery* 2015;**33**(2):61-8.

Paula 2011 *{published data only}*

De Paula AB, Campos JADB, Diniz MB, Hebling J, Rodrigues JA. In situ and in vitro comparison of laser fluorescence with visual inspection in detecting occlusal caries lesions. *Lasers in Medical Science* 2011;**26**:1-5.

Pereira 2011 *{published data only}*

Pereira AC, Eggertsson H, González-Cabezas C, Zero DT, Eckert GJ, Mialhe FL. Quantitative light-induced fluorescence (QLF) in relation to other technologies and conventional methods for detecting occlusal caries in permanent teeth. *Brazilian Journal of Oral Sciences* 2011;**10**(1):27-32.

Pinelli 2002 *{published data only}*

Pinelli C, Campos Serra M, de Castro Monteiro Loffredo L. Validity and reproducibility of a laser fluorescence system for detecting the activity of white-spot lesions on free smooth surfaces in vivo. *Caries Research* 2002;**36**(1):19-24.

Pourhashemi 2009 *{published data only}*

Pourhashemi SJ, Jafari A, Motahhari P, Panjnoosh M, Kharrazi Fard MJ, Sanati I, et al. An in-vitro comparison of visual inspection, bite-wing radiography, and laser fluorescence methods for the diagnosis of occlusal caries. *Journal of the Indian Society of Pedodontics and Preventive Dentistry* 2009;**27**(2):90-3.

Presoto 2017 *{published data only}*

Presoto CD, Trevisan TC, Andrade MC, Dantas AA, Campos JADB, Oliveira-Junior OB. Clinical effectiveness of fluorescence, digital images and ICDAS for detecting occlusal caries. *Revista de Odontologia da UNESP* 2017;**46**(2):109-15.

Rando-Meirelles 2011 *{published data only}*

Rando-Meirelles MP, de Sousa Mda L. Using laser fluorescence (DIAGNOdent) in surveys for the detection of noncavitated occlusal dentine caries. *Community Dental Health* 2011;**28**(1):17-21.

Reis 2004 *{published data only}*

Reis A, Zach VL Jr, de Lima AC, de Lima Navarro MF, Grande RH. Occlusal caries detection: a comparison of DIAGNOdent and two conventional diagnostic methods. *Journal of Clinical Dentistry* 2004;**15**(3):76-82.

Reis 2006 *{published data only}*

Reis A, Mendes FM, Angnes V, Angnes G, Grande RH, Loguercio AD. Performance of methods of occlusal caries detection in permanent teeth under clinical and laboratory conditions. *Journal of Dentistry* 2006;**34**(2):89-96.

Ribeiro 2015 *{published data only}*

Ribeiro AA, Purger F, Rodrigues JA, Oliveira PR, Lussi A, Monteiro AH, et al. Influence of contact points on the performance of caries detection methods in approximal surfaces of primary molars: an in vivo study. *Caries Research* 2015;**49**(2):99-108.

Rocha 2003 *{published data only}*

Rocha RO, Ardenghi TM, Oliveira LB, Rodrigues CR, Ciamponi AL. In vivo effectiveness of laser fluorescence compared to visual inspection and radiography for the detection of occlusal caries in primary teeth. *Caries Research* 2003;**37**(6):437-41.

Rocha-Cabral 2008 *{published data only}*

Rocha-Cabral RM, Mendes FM, Miura F, Ribeiro Ada C, Braga MM, Zezell DM. Autoclaving and battery capacity influence on laser fluorescence measurements. *Acta Odontologica Scandinavica* 2008;**66**(2):122-7.

Rodrigues 2008 *{published data only}*

Rodrigues JA, Hug I, Diniz MB, Lussi A. Performance of fluorescence methods, radiographic examination and ICDAS II on occlusal surfaces in vitro. *Caries Research* 2008;**42**(4):297-304.

Rodrigues 2009 *{published data only}*

Rodrigues JA, Diniz MB, Josgrilberg EB, Cordeiro RC. In vitro comparison of laser fluorescence performance with visual examination for detection of occlusal caries in permanent and primary molars. *Lasers in Medical Science* 2009;**24**(4):501-6.

Rodrigues 2011 *{published data only}*

Rodrigues JA, Hug I, Neuhaus KW, Lussi A. Light-emitting diode and laser fluorescence-based devices in detecting occlusal caries. *Journal of Biomedical Optics* 2011;**16**(10):107003.

Seremidi 2012 *{published data only}*

Seremidi K, Lagouvardos P, Kavvadia K. Comparative in vitro validation of VistaProof and DIAGNOdent pen for occlusal caries detection in permanent teeth. *Operative Dentistry* 2012;**37**(3):234-45.

Sheehy 2001 *{published data only}*

Sheehy EC, Brailsford SR, Kidd EA, Beighton D, Zoitopoulos L. Comparison between visual examination and a laser fluorescence system for in vivo diagnosis of occlusal caries. *Caries Research* 2001;**35**:421-6.

Shi 2000 *{published data only}*

Shi XQ, Welander U, Angmar-Mansson B. Occlusal caries detection with KaVo DIAGNOdent and radiography: an in vitro comparison. *Caries Research* 2000;**34**:151-8.

Shwetha 2017 *{published data only}*

Shwetha G, Chandra P, Anandakrishna L, Dhananjaya G, Shetty AK, Kamath PS. Validation of different diagnostic aids in detection of occlusal caries in primary molars: an in vitro study. *Journal of the Indian Society of Pedodontics and Preventive Dentistry* 2017;**35**(4):301-6.

Sinanoglu 2014 *{published data only}*

Sinanoglu A, Ozturk E, Ozel E. Diagnosis of occlusal caries using laser fluorescence versus conventional methods in permanent posterior teeth: a clinical study. *Photomedicine and Laser Surgery* 2014;**32**(3):130-7.

Souza 2013 *{published data only}*

Souza JF, Boldieri T, Diniz MB, Rodrigues JA, Lussi A, Cordeiro RC. Traditional and novel methods for occlusal caries detection: performance on primary teeth. *Lasers in Medical Science* 2013;**28**:287-95.

Souza 2014 *{published data only}*

de Souza JF, Diniz MB, Boldieri T, Rodrigues JA, Lussi A, de Cassia Loiola Cordeiro R. In vitro performance of a pen-type laser fluorescence device and bitewing radiographs for approximal caries detection in permanent and primary teeth. *Indian Journal of Dental Research* 2014;**25**(6):702-10.

Souza 2018 *{published data only}*

de Souza LA, Cancio V, Tostes MA. Accuracy of pen-type laser fluorescence device and radiographic methods in detecting approximal carious lesions in primary teeth - an in vivo study. *International Journal of Paediatric Dentistry* 2018;**28**(5):472-80.

Sridhar 2009 *{published data only}*

Sridhar N, Tandon S, Rao N. A comparative evaluation of DIAGNOdent with visual and radiography for detection of occlusal caries: an in vitro study. *Indian Journal of Dental Research* 2009;**20**:326-31.

Teo 2014 *{published data only}*

Teo TK, Ashley PF, Louca C. An in vivo and in vitro investigation of the use of ICDAS, DIAGNOdent pen and CarieScan PRO for the detection and assessment of occlusal caries in primary molar teeth. *Clinical Oral Investigations* 2014;**18**(3):737-44.

Tonioli 2002 *{published data only}*

Tonioli MB, Bouschlicher MR, Hillis SL. Laser fluorescence detection of occlusal caries. *American Journal of Dentistry* 2002;**15**:268-73.

Tonkaboni 2018 *{published data only}*

Tonkaboni A, Saffarpour A, Aghapourzangeneh F, Fard MJK. Comparison of diagnostic effects of infrared imaging and bitewing radiography in proximal caries of permanent teeth. *Lasers in Medical Science* 2018;**34**(5):873-9.

Umemori 2010 *{published data only}*

Umemori S, Tonami K, Nitta H, Mataki S, Araki K. The possibility of digital imaging in the diagnosis of occlusal caries. *International Journal of Dentistry* 2010;**2010**:860515.

Valera 2008 *{published data only}*

Valera FB, Pessan JP, Valera R, Mondelli J, Percinoto C. Comparison of visual inspection, radiographic examination, laser fluorescence and their combinations on treatment decisions for occlusal surfaces. *American Journal of Dentistry* 2008;**21**:25-9.

Van Hilsen 2013 *{published data only}*

Van Hilsen Z, Jones RS. Comparing potential early caries assessment methods for teledentistry. *BMC Oral Health* 2013;**13**:16.

Virajsilp 2005 *{published data only}*

Virajsilp V, Thearmontree A, Aryatawong S, Paiboonwarachat D. Comparison of proximal caries detection in primary teeth between laser fluorescence and bitewing radiography. *Pediatric Dentistry* 2005;**27**:493-9.

Yoon 2017 *{published data only}*

Yoon HI, Yoo MJ, Park EJ. Detection of proximal caries using quantitative light-induced fluorescence-digital and laser fluorescence: a comparative study. *Journal of Advanced Prosthodontics* 2017;**9**(6):432-8.

Zeitouny 2014 *{published data only}*

Zeitouny M, Feghali M, Nasr A, Abou-Samra P, Saleh N, Bourgeois D, et al. SOPROLIFE system: an accurate diagnostic enhancer. *Scientific World Journal* 2014;**2014**:924741.

References to studies excluded from this review

Abalos 2009 *{published data only}*

Abalos C, Herrera M, Jimenez-Planas A, Llamas R. Performance of laser fluorescence for detection of occlusal dentinal caries lesions in permanent molars: an in vivo study with total validation of the sample. *Caries Research* 2009;**43**:137-41.

Abalos 2012 *{published data only}*

Abalos C, Mendoza A, Jimenez-Planas A, Guerrero E, Chaparro A, Garcia-Godoy F. Performance of laser fluorescence for the detection of enamel caries in non-cavitated occlusal surfaces: clinical study with total validation of the sample. *American Journal of Dentistry* 2012;**25**:44-8.

Abou 2016 *{published data only}*

Abou Nader C, Pellen F, Loutfi H, Mansour R, Le Jeune B, Le Brun G, et al. Early diagnosis of teeth erosion using polarized laser speckle imaging. *Journal of Biomedical Optics* 2016;**21**:71103.

Abrams 2017 *{published data only}*

Abrams TE, Abrams SH, Sivagurunathan KS, Silvertown JD, Hellen WMP, Elman GI, et al. In vitro detection of caries around amalgam restorations using four different modalities. *Open Dentistry Journal* 2017;**11**:609.

Amaechi 2013 *{published data only}*

Amaechi BT, Chedjieu I, Lozano-Pineda J. Clinical evaluation of an enhanced white light and fluorescence device for early detection of caries lesions. *Journal of Clinical Dentistry* 2013;**24**:43-8.

Anttonen 2004 *{published data only}*

Anttonen V, Seppa L, Hausen H. A follow-up study of the use of DIAGNOdent for monitoring fissure caries in children. *Community Dentistry and Oral Epidemiology* 2004;**32**:312-8.

Askaroglou 2011 *{published data only}*

Askaroglou E, Kavvadia K, Lagouvardos P, Papagiannoulis L. Effect of sealants on laser fluorescence caries detection in primary teeth. *Lasers in Medical Science* 2011;**26**:29-34.

Betrisey 2014 *{published data only}*

Betrisey E, Rizcalla N, Krejci I, Ardu S. Caries diagnosis using light fluorescence devices: vistaProof and DIAGNOdent. *Odontology* 2014;**102**(2):330-5.

Blazejewska 2016 *{published data only}*

Blazejewska A, Dacyna N, Niesiobedzki P, Trzaska M, Gozdowski D, Turska-Szybka A, et al. Comparison of the detection of proximal caries in children and youth using DIAGNOcam and bitewing radiovisiography. *Dental and Medical Problems* 2016;**53**:468-75.

Diniz 2016 *{published data only}*

Diniz MB, Cordeiro RC, Ferreira-Zandona AG. Detection of caries around amalgam restorations on approximal surfaces. *Operative Dentistry* 2016;**41**:34-43.

Gomez 2013 *{published data only}*

Gomez J, Zakian C, Salsone S, Pinto SC, Taylor A, Pretty IA, et al. In vitro performance of different methods in detecting occlusal caries lesions. *Journal of Dentistry* 2013;**41**(2):180-6.

Heinrich-Weltzien 2005 *{published data only}*

Heinrich-Weltzien R, Kuhnisch J, Ifland S, Tranaeus S, Angmar-Mansson B, Stosser L. Detection of initial caries lesions on smooth surfaces by quantitative light-induced fluorescence and visual examination: an in vivo comparison. *European Journal of Oral Sciences* 2005;**113**(6):494-8.

Holtzman 2014 *{published data only}*

Holtzman JS, Ballantine J, Fontana M, Wang A, Calantog A, Benavides E, et al. Assessment of early occlusal caries pre- and post-sealant application - an imaging approach. *Lasers in Surgery and Medicine* 2014;**46**(6):499-507.

Jablonski-Momeni 2011a *{published data only}*

Jablonski-Momeni A, Ricketts DN, Rolfsen S, Stoll R, Heinzel-Gutenbrunner M, Stachniss V, et al. Performance of laser fluorescence at tooth surface and histological section. *Lasers in Medical Science* 2011;**26**:171-8.

Jablonski-Momeni 2013 *{published data only}*

Jablonski-Momeni A, Liebegall F, Stoll R, Heinzel-Gutenbrunner M, Pieper K. Performance of a new fluorescence camera for detection of occlusal caries in vitro. *Lasers in Medical Science* 2013;**28**(1):101-9.

Jallad 2015 *{published data only}*

Jallad M, Zero D, Eckert G, Ferreira Zandona A. In vitro detection of occlusal caries on permanent teeth by a visual, light-induced fluorescence and photothermal radiometry and modulated luminescence methods. *Caries Research* 2015;**49**(5):523-30.

Kordic 2003 *{published data only}*

Kordic A, Lussi A, Luder HU. Performance of visual inspection, electrical conductance and laser fluorescence in detecting occlusal caries in vitro. *Schweizer Monatsschrift fur Zahnmedizin* 2003;**113**(8):852-9.

Marinova-Takorova 2014 *{published data only}*

Marinova-Takorova M, Anastasova R, Panov VE. Comparative evaluation of the effectiveness of five methods for early diagnosis of occlusal caries lesions - in vitro study. *Journal of IMAB* 2014;**20**(3):533-6.

Melo 2015 *{published data only}*

Melo M, Pascual A, Camps I, Del Campo A. In vivo study of different methods for diagnosing pit and fissure caries. *Journal of Clinical and Experimental Dentistry* 2015;**7**(3):e387-91.

Menem 2017 *{published data only}*

Menem R, Barngkgei I, Beiruti N, Al Haffar I, Joury E. The diagnostic accuracy of a laser fluorescence device and digital radiography in detecting approximal caries lesions in posterior permanent teeth: an in vivo study. *Lasers in Medical Science* 2017;**32**(3):621-8.

Mujat 2003 *{published data only}*

Mujat C, van der Veen MH, Ruben JL, ten Bosch JJ, Dogariu A. Optical path-length spectroscopy of incipient caries lesions in relation to quantitative light-induced fluorescence and lesion characteristics. *Applied Optics* 2003;**42**(16):2979-86.

Mujat 2004 *{published data only}*

Mujat C, Van Der Veen MH, Ruben JL, Dogariu A, Ten Bosch JJ. The influence of drying on quantitative laser fluorescence and optical path lengths in incipient natural caries lesions. *Caries Research* 2004;**38**(5):484-92.

Nemes 2001 *{published data only}*

Nemes J, Csillag M, Toth Z, Fazekas A. Reproducibility of the laser fluorescence method for the diagnosis of occlusal caries. Clinical study. *Fogorvosi Szemle* 2001;**94**(1):33-6.

Parviainen 2013 *{published data only}*

Parviainen H, Vahanikkila H, Laitala ML, Tjaderhane L, Anttonen V. Evaluating performance of dental caries detection methods among third-year dental students. *BMC Oral Health* 2013;**13**:70.

Patel 2014 *{published data only}*

Patel SA, Shepard WD, Barros JA, Streckfus CF, Quock RL. In vitro evaluation of Midwest Caries ID: a novel light-emitting diode for caries detection. *Operative Dentistry* 2014;**39**(6):644-51.

Pereira 2009 *{published data only}*

Pereira AC, Eggertsson H, Martinez-Mier EA, Mialhe FL, Eckert GJ, Zero DT. Validity of caries detection on occlusal surfaces and treatment decisions based on results from multiple caries-detection methods. *European Journal of Oral Sciences* 2009;**117**(1):51-7.

Rechmann 2012 *{published data only}*

Rechmann P, Charland D, Rechmann BM, Featherstone JD. Performance of laser fluorescence devices and visual examination for the detection of occlusal caries in permanent molars. *Journal of Biomedical Optics* 2012;**17**(3):036006.

Subka 2019 *{published data only}*

Subka S, Rodd H, Nugent Z, Deery C. In vivo validity of proximal caries detection in primary teeth, with histological validation. *International Journal of Paediatric Dentistry* 2019;**29**(4):429-38.

Theocharopoulou 2015 *{published data only}*

Theocharopoulou A, Lagerweij MD, van Strijp AJ. Use of the ICDAS system and two fluorescence-based intraoral devices for examination of occlusal surfaces. *European Journal of Paediatric Dentistry* 2015;**16**(1):51-5.

Zhang 2009 *{published data only}*

Zhang W, McGrath C, Lo ECM. A comparison of root caries diagnosis based on visual-tactile criteria and DIAGNOdent in vivo. *Journal of Dentistry* 2009;**37**(7):509-13.

Additional references

Amaechi 2019

Amaechi Bennett T. Photothermal radiometry and modulated luminescence: the Canary system. In: Detection and Assessment of Dental Caries. Springer, 2019:177-86.

Anaise 1984

Anaise JZ. Measurement of dental caries experience - modification of the DMFT index. *Community Dentistry and Oral Epidemiology* 1984;**12**(1):43-6.

Angmar-Månsson 2001

Angmar-Månsson B, Ten Bosch JJ. Quantitative light-induced fluorescence (QLF): a method for assessment of incipient caries lesions. *Dento Maxillo Facial Radiology* 2001;**30**:298-307.

Backer Dirks 1951

Backer Dirks O, Van Amerongen J, Winkler KC. A reproducible method for caries evaluation. *Journal of Dental Research* 1951;**30**:346-59.

Bader 2002

Bader JD, Shugars DA, Bonito AJ. A systematic review of the performance of methods for identifying carious lesions. *Journal of Public Health Dentistry* 2002;**62**(4):201-13.

Bader 2004

Bader JD, Shugars DA. A systematic review of the performance of a laser fluorescence device for detecting caries. *Journal of the American Dental Association* 2004;**135**(10):1413-26.

Bakhshandeh 2018

Bakhshandeh A, Floriano I, Braga MM, Thorlacius KA, Ekstrand KR. Relationship between depth of approximal caries lesions and presence of bacteria in the dentine in primary and permanent posterior teeth: a radiographic examination with microbiological evaluation. *Acta Odontologica Scandinavica* 2018;**76**(7):509-14.

Bloemendal 2004

Bloemendal E, De Vet HC, Bouter LM. The value of bitewing radiographs in epidemiological caries research: a systematic review of the literature. *Journal of Dentistry* 2004;**32**:255-64.

Bossuyt 2003

Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, et al. Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. *Radiology* 2003;**226**(1):24-8.

Bossuyt 2015

Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig L, et al. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. *Clinical Chemistry* 2015;**61**(12):1446-52.

Broadbent 2008

Broadbent JM, Thomson WM, Poulton R. Trajectory patterns of dental caries experience in the permanent dentition to the fourth decade of life. *Journal of Dental Research* 2008;**87**:69-72.

Children's Dental Health Survey 2013

Pitts N, Chadwick B, Anderson T, Ramsay G. Children's Dental Health Survey 2013. Report 2: Dental Disease and Damage in Children England, Wales and Northern Ireland. files.digital.nhs.uk/publicationimport/pub17xxx/pub17137/cdhs2013-report2-dental-disease.pdf 2015.

Deeks 2005

Deeks JJ, Macaskill P, Irwig L. The performance of tests of publication bias and other sample size effects in systematic reviews of diagnostic test accuracy was assessed. *Journal of Clinical Epidemiology* 2005;**58**(9):882-93.

Deeks 2013

Deeks JJ, Bossuyt PM, Gatsonis C, editor(s). Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 1.0. The Cochrane Collaboration, 2013. Available from srdta.cochrane.org.

de Vet 2008

de Vet HCW, Eisinga A, Riphagen II, Aertgeerts B, Pewsner D. Chapter 7: Searching for studies. In: Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 0.4 (updated September 2008). The Cochrane Collaboration, 2008. Available from srdta.cochrane.org.

Dinnes 2016

Dinnes J, Mallett S, Hopewell S, Roderick PJ, Deeks JJ. The Moses–Littenberg meta-analytical method generates systematic differences in test accuracy compared to hierarchical meta-analytical models. *Journal of Clinical Epidemiology* 2016;**80**:77-87.

Downer 1975

Downer MC. Concurrent validity of an epidemiological diagnostic system for caries with the histological appearance of extracted teeth as validating criterion. *Caries Research* 1975;**9**:231-46.

Dye 2015

Dye BA, Thornton-Evans G, Li X, Iafolla TJ. Dental caries and sealant prevalence in children and adolescents in the United States, 2011-2012. *NCHS Data Brief* 2015;**191**:1-8.

Ekstrand 1997

Ekstrand KR, Ricketts DNJ, Kidd EAM. Reproducibility and accuracy of three methods for assessment of demineralization

depth on the occlusal surface: an in vitro examination. *Caries Research* 1997;**31**:224-31.

Ekstrand 1998

Ekstrand KR, Ricketts DNJ, Kidd EAM, Qvist V, Schou S. Detection, diagnosing, monitoring and logical treatment of occlusal caries in relation to lesion activity and severity: an in vivo examination with histological validation. *Caries Research* 1998;**32**:247-54.

Ekstrand 2007

Ekstrand KR, Martignon S, Ricketts DJN, Qvist V. Detection and activity assessment of primary coronal caries lesions: a methodologic study. *Operative Dentistry* 2007;**32**(3):225-35.

Ellwood 1997

Ellwood RP, Davies RM, Worthington HV. Evaluation of a dental subtraction radiography system. *Journal of Periodontal Research* 1997;**32**:241-8.

Featherstone 2004

Featherstone JDB. The continuum of dental caries - evidence for a dynamic disease process. *Journal of Dental Research* 2004;**83**:39-42.

Feigin 2016

Feigin V. Global, regional, and national incidence, prevalence, and years lived with disability for 310 acute and chronic diseases and injuries, 1990-2015: a systematic analysis for the global burden of disease study 2015. *Lancet* 2016;**388**:1545-602.

Fejerskov 2015

Fejerskov O, Nyvad B, Kidd E, editor(s). Dental Caries: The Disease and its Clinical Management. 3rd edition. Wiley-Blackwell, 2015.

Fyffe 2000

Fyffe HE, Deery C, Nugent ZJ, Nuttall NM, Pitts NB. Effect of diagnostic threshold on the validity and reliability of epidemiological caries diagnosis using the Dundee Selectable Threshold Method for caries diagnosis (DSTM). *Community Dentistry and Oral Epidemiology* 2000;**28**:42-51.

Gimenez 2013

Gimenez T, Braga MM, Raggio DP, Deery C, Ricketts DN, Mendes FM. Fluorescence-based methods for detecting caries lesions: systematic review, meta-analysis and sources of heterogeneity. *PLOS One* 2013;**8**(4):e60421.

Gopalakrishna 2014

Gopalakrishna G, Mustafa RA, Davenport C, Scholten RJPM, Hyde C, Brozek J, et al. Applying Grading of Recommendations Assessment, Development and Evaluation (GRADE) to diagnostic tests was challenging but doable. *Journal of Clinical Epidemiology* 2014;**67**:760-8.

Hall-Scullin 2017

Hall-Scullin E, Whitehead H, Milsom K, Tickle M, Su T-L, Walsh T. Longitudinal study of caries development from childhood to adolescence. *Journal of Dental Research* 2017;**96**:762-7.

Horner 2009

Horner K, Islam M, Flygare L, Tsiklakis K, Whaites E. Basic principles for use of dental cone beam computed tomography: consensus guidelines of the European Academy of Dental and Maxillofacial Radiology. *Dento Maxillo Facial Radiology* 2009;**38**:187-95.

Hse 2015

Hse K. The challenge of oral disease - a call for global action. In: The Oral Health Atlas. 2nd edition. Geneva: FDI World Dental Federation, 2015:8-11.

Hsu 2011

Hsu J, Brożek JL, Terracciano L, Kreis J, Compalati E, Stein AT, et al. Application of GRADE: making evidence-based recommendations about diagnostic tests in clinical practice guidelines. *Implementation Science* 2011;**6**(1):62.

Ismail 2004

Ismail AI. Visual and visuo-tactile detection of dental caries. *Journal of Dental Research* 2004;**83**:56-66.

Ismail 2007

Ismail AI, Sohn W, Tellez M, Amaya A, Sen A, Hasson H, et al. The International Caries Detection and Assessment System (ICDAS): an integrated system for measuring dental caries. *Community Dentistry and Oral Epidemiology* 2007;**35**:170-8.

Ismail 2013

Ismail AI, Tellez M, Pitts NB, Ekstrand KR, Ricketts D, Longbottom C, et al. Caries management pathways preserve dental tissues and promote oral health. *Community Dentistry and Oral Epidemiology* 2013;**41**(1):e12-40.

Ismail 2015

Ismail AI, Pitts NB, Tellez M. The International Caries Classification and Management System (ICCMS™) an example of a caries management pathway. *BMC Oral Health* 2015;**15 Suppl 1**:S9. [DOI: 10.1186/1472-6831-15-S1-S9]

Jones 2017

Jones CM, Davies GM, Monaghan N, Morgan MZ, Neville JS, Pitts NB. The caries experience of 5 year-old children in Scotland in 2013-2014, and in England and Wales in 2014-2015. Reports of cross-sectional dental surveys using BASCD criteria. *Community Dental Health* 2017;**34**:157-62.

Kassebaum 2015

Kassebaum NJ, Bernabé E, Dahiya M, Bhandari B, Murray CJL, Marcenes W. Global burden of untreated caries: a systematic review and metaregression. *Journal of Dental Research* 2015;**94**:650-8.

Kidd 2004

Kidd EAM, Fejerskov O. What constitutes dental caries? Histopathology of carious enamel and dentin related to the action of cariogenic biofilms. *Journal of Dental Research* 2004;**83**:35-8.

Kidd 2016

Kidd EAM, Fejerskov O. Essentials of Dental Caries. 4th edition. Oxford University Press, 2016.

Kim 2019

Kim Baek-Il. Quantitative light-induced fluorescence. In: Detection and Assessment of Dental Caries. Springer, 2019:159-70.

Leeflang 2008

Leeflang MMG, Deeks JJ, Gatsonis C, Bossuyt PMM. Systematic reviews of diagnostic test accuracy. *Annals of Internal Medicine* 2008;**149**:889-97.

Listl 2015

Listl S, Galloway J, Mossey PA, Marcenes W. Global economic impact of dental diseases. *Journal of Dental Research* 2015;**94**:1355-61.

Macaskill 2010

Macaskill P, Gatsonis C, Deeks JJ, Harbord RM, Takwoingi Y. Chapter 10: Analysing and presenting results. In: Deeks JJ, Bossuyt PM, Gatsonis C, editor(s). Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 1.0. The Cochrane Collaboration, 2010. Available from srdta.cochrane.org.

Macey 2018

Macey R, Walsh T, Riley P, Glenny AM, Worthington HV, Clarkson JE, et al. Tests to detect and inform the diagnosis of caries. *Cochrane Database of Systematic Reviews* 2018, Issue 12. Art. No: CD013215. [DOI: 10.1002/14651858.CD013215]

McInnes 2018

McInnes MDF, Moher D, Thombs BD, McGrath TA, Bossuyt PM, Clifford T, et al. Preferred reporting items for a systematic review and meta-analysis of diagnostic test accuracy studies: the PRISMA-DTA statement. *JAMA* 2018;**319**:388-96.

Milsom 2003

Milsom KM, Tickle M, Humphris GM, Blinkhorn AS. The relationship between anxiety and dental treatment experience in 5-year-old children. *British Dental Journal* 2003;**194**:503.

Moher 2009

Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Annals of Internal Medicine* 2009;**151**(4):264-9.

Neuhaus 2016

Neuhaus KW, Lussi A. New caries diagnostic methods. In: Understanding Dental Caries. Springer, 2016:53-61.

Neuhaus 2019

Neuhaus KW, Lussi A. DIAGNOdent. In: Detection and Assessment of Dental Caries. Springer, 2019:171-5.

Nyvad 1997

Nyvad B, Fejerskov O. Assessing the stage of caries lesion activity on the basis of clinical and microbiological examination. *Community Dentistry and Oral Epidemiology* 1997;**25**:69-75.

Nyvad 1999

Nyvad B, Machiulskiene V, Bælum V. Reliability of a new caries diagnostic system differentiating between active and inactive caries lesions. *Caries Research* 1999;**33**:252-60.

Petersen 2005

Petersen PE, Bourgeois D, Ogawa H, Estupinan-Day S, Ndiaye C. The global burden of oral diseases and risks to oral health. *Bulletin of the World Health Organization* 2005;**83**:661-9.

Pitts 1988

Pitts NB, Fyffe HE. The effect of varying diagnostic thresholds upon clinical caries data for a low prevalence group. *Journal of Dental Research* 1988;**67**:592-6.

Pitts 1997

Pitts NB, Evans DJ, Pine CM. British Association for the Study of Community Dentistry (BASCD) diagnostic criteria for caries prevalence surveys-1996/97. *Community Dental Health* 1997;**14**:6-9.

Pitts 2001

Pitts NB. Clinical diagnosis of dental caries: a European perspective. *Journal of Dental Education* 2001;**65**:972-8.

Pitts 2009

Pitts N. Detection, Assessment, Diagnosis and Monitoring of Caries. Switzerland: Karger Basel, 2009.

Pitts 2017

Pitts NB, Zero DT, Marsh PD, Ekstrand K, Weintraub JA, Ramos-Gomez F, et al. Dental caries. *Nature Reviews Disease Primers* 2017;**3**:17030.

Pretty 2006

Pretty IA. Caries detection and diagnosis: novel technologies. *Journal of Dentistry* 2006;**34**:727-39.

Public Health England 2014

Public Health England. Oral health survey of five-year-old. Dental public health epidemiology programme. www.nwph.net/dentalhealth/14_15_5yearold/Protocol_2014_15_5%20yr%20olds%20v2.pdf (accessed 15 May 2018).

Reitsma 2005

Reitsma JB, Glas AS, Rutjes AWS, Scholten RJPM, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *Journal of Clinical Epidemiology* 2005;**58**:982-90.

Review Manager 2020 [Computer program]

Nordic Cochrane Centre, The Cochrane Collaboration Review Manager 5 (RevMan 5). Version 5.4. Copenhagen: Nordic Cochrane Centre, The Cochrane Collaboration, 2020.

Rutter 2001

Rutter CM, Gatsonis CA. A hierarchical regression approach to meta-analysis of diagnostic test accuracy evaluations. *Statistics in Medicine* 2001;**20**(19):2865-84.

Schünemann 2020

Schünemann HJ, Mustafa RA, Brozek J, Steingart KR, Leeflang M, Murad MH, et al. GRADE guidelines: 21 part 2. Test accuracy: inconsistency, imprecision, publication bias, and other domains for rating the certainty of evidence and presenting it in evidence profiles and summary of findings tables. *Journal of Clinical Epidemiology* 2020;**122**:142-52.

Schünemann 2020a

Schünemann HJ, Mustafa RA, Brozek J, Steingart KR, Leeflang M, Murad MH, et al. GRADE guidelines: 21 part 1. Study design, risk of bias, and indirectness in rating the certainty across a body of evidence for test accuracy. *Journal of Clinical Epidemiology* 2020;**122**:129-41.

Shoaib 2009

Shoaib L, Deery CM, Ricketts DNJ, Nugent ZJ. Validity and reproducibility of ICDAS II in primary teeth. *Caries Research* 2009;**43**:442-8.

Steele 2011

Steele J, O' Sullivan I. Executive summary: Adult Dental Health Survey 2009. files.digital.nhs.uk/publicationimport/pub01xxx/pub01086/adul-dent-heal-surv-summ-them-exec-2009-rep2.pdf 2011.

Takwoingi 2010

Takwoingi Y, Deeks JJ. MetaDAS: a SAS macro for meta-analysis of diagnostic accuracy studies. Quick reference and worked example. srdta.cochrane.org/ 2010.

Takwoingi 2015

Takwoingi Y, Riley RD, Deeks JJ. Meta-analysis of diagnostic accuracy studies in mental health. *Evidence-Based Mental Health* 2015;**18**:103-9.

Tam 2001

Tam LE, McComb D. Diagnosis of occlusal caries: Part II. Recent diagnostic technologies. *Journal of the Canadian Dental Association* 2001;**67**:459-64.

Thomson 2000

Thomson WM, Locker D, Poulton R. Incidence of dental anxiety in young adults in relation to dental treatment experience. *Community Dentistry and Oral Epidemiology* 2000;**28**:289-94.

Wenzel 2006

Wenzel A, Anthonisen PN, Juul MB. Reproducibility in the assessment of caries lesion behaviour: a comparison between conventional film and subtraction radiography. *Caries Research* 2000;**34**:214-8.

White 2012

White DA, Tsakos G, Pitts NB, Fuller E, Douglas GVA, Murray JJ, et al. Adult Dental Health Survey 2009: common oral health conditions and their impact on the population. *British Dental Journal* 2012;**213**(11):567.

Whiting 2011

Whiting PF, Rutjes AWS, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Annals of Internal Medicine* 2011;**155**:529-36.

Wilson 1968

Wilson JMG, Jungner G, World Health Organization. Principles and Practice of Screening for Disease. Geneva: World Health Organization, 1968.

World Health Organization 2017

World Health Organization. Sugars and dental caries. WHO Technical Information Note. apps.who.int/iris/bitstream/handle/10665/259413/WHO-NMH-NHD-17.12-eng.pdf?sequence=1 (accessed 15 May 2018).

Young 2015

Young DA, Nový BB, Zeller GG, Hale R, Hart TC, Truelove EL. The American Dental Association Caries Classification System for clinical practice: a report of the American Dental Association Council on Scientific Affairs. *Journal of the American Dental Association* 2015;**146**(2):79-86.

Zhang 2019

Zhang Y, Coello PA, Guyatt GH, Yepes-Nuñez JJ, Akl EA, Hazlewood G, et al. GRADE guidelines: 20. Assessing the certainty of evidence in the importance of outcomes or values and preferences - inconsistency, imprecision, and other domains. *Journal of Clinical Epidemiology* 2019;**111**:83-93.

CHARACTERISTICS OF STUDIES

Characteristics of included studies *[ordered by study ID]*

Achilleos 2013

Study characteristics	
Patient Sampling	Method of sampling: selected
	Included conditions: no cavitation
	Teeth: permanent molars and premolars

Achilleos 2013 *(Continued)*

	Sealants: unclear Surface: occlusal
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported Country: Greece Setting: extracted for orthodontic purposes Number of participants/teeth/sites: 38 teeth Prevalence: enamel 0.95, dentine 0.39
Index tests	Category of test: DIAGNOdent pen and VistaProof Sequence of test(s): visual, then index tests, then reference standard Examiner training and calibration: experienced, trained, and calibrated dentists Teeth cleaning prior to examination: calculus and debris were removed by paste and brush burr Tooth drying prior to examination: yes Threshold applied: DIAGNOdent pen: 0-13 sound, 14-20 enamel (outer), 21-29 enamel (deep), > 30 dentinal VistaProof: "software shows the region of the teeth that emits fluorescence and an outcome value in different colors, ranging from 0 to 5, which defines the caries lesions extension according to the manufacturer's recommendations. Numerical and color scales were: 1.0–1.5/blue shows beginning enamel caries, 1.5–2.0/red shows deep enamel caries, 2.0–2.5/orange shows dentin caries, and 2.5–5.0/yellow shows deep dentin caries" Device specifics: sapphire fibre tip
Target condition and reference standard(s)	Category: histology Sequence of index test and reference standard: following index test Training of examiner: experienced, same examiner as index test Blinding to index test: no Multiple tests: no Site selection: 3 sections Target condition: caries free, early enamel, deep enamel, outer dentine, dentine, deep dentine
Flow and timing	Participants with index test but no reference standard: 0

Achilleos 2013 *(Continued)*

	Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0		
Comparative			
Notes	Multiple examiners reported so examiner one values reported		
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	

Achilleos 2013 *(Continued)*

Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Red fluorescence)			
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Akarsu 2006

Study characteristics	
Patient Sampling	Method of sampling: unclear Included conditions: unclear, "suspected to have occlusal caries" but unclear to what level Teeth: permanent molars (third molars excluded) Sealants: no Surface: occlusal
Patient characteristics and setting	Age: 18 to 25 years Sex: 87 female, 74 male Ethnicity: not reported Country: Turkey Setting: restorative clinic at dental hospital Number of participants/teeth/sites: 161 participants, 187 teeth Prevalence: enamel 0.77, dentine 0.52

Akarsu 2006 *(Continued)*

Index tests	Category of test: DIAGNOdent Sequence of test(s): visual, radiograph, DIAGNOdent, then reference standard (visual, radiograph, and DIAGNOdent used as part of reference standard) Examiner training and calibration: unclear Teeth cleaning prior to examination: calculus and plaque removed using a scaler and rubber cup - no pumice used Tooth drying prior to examination: 8 seconds Threshold applied: calculated in study: 0-5.5 sound, 5.5-11.5 enamel, 11.5 superficial dentine, 18.5+ deep dentine Device specifics: probe A, conical tip
Target condition and reference standard(s)	Category: teeth identified as carious by the index tests were "removed by using rotational cutting devices" and the cavities assessed visually, i.e. excavation Sequence of index test and reference standard: following index test Training of examiner: experienced, same examiner as index test Blinding to index test: no Multiple tests: no Site selection: 3 sections Target condition: caries free, early enamel, deep enamel, outer dentine, dentine, deep dentine
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0
Comparative	
Notes	First observer results used

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		

Akarsu 2006 *(Continued)*

Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	No		
If a threshold was used, was it pre-specified?	No		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	No		
If a threshold was used, was it pre-specified?	No		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	No		
Were the reference standard results interpreted without knowledge of the results of the index tests?	No		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			

Akarsu 2006 *(Continued)*

Was there an appropriate interval between index test and reference standard?	Yes	
Did all patients receive the same reference standard?	No	
Were all patients included in the analysis?	Yes	
Could the patient flow have introduced bias?		High risk

Aktan 2012

Study characteristics	
Patient Sampling	Method of sampling: selected Included conditions: no cavitation and enamel lesions Teeth: permanent molars Sealants: unclear Surface: occlusal
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported Country: Turkey Setting: extracted teeth Number of participants/teeth/sites: 83 teeth/129 sites Prevalence: enamel 0.58, dentine 0.21
Index tests	Category of test: DIAGNOdent pen and Midwest Sequence of test(s): before reference standard Examiner training and calibration: yes Teeth cleaning prior to examination: calculus removed Tooth drying prior to examination: not reported Threshold applied: DIAGNOdent pen: 0-13 sound, 14-20 enamel, > 20 dentine Midwest: manufacturer recommendations; no signal/green light - sound; slow or medium signal/red light - enamel; rapid or continuous signal/red light - dentine Device specifics: DIAGNOdent pen cylindrical tip
Target condition and reference standard(s)	Category: histology Sequence of index test and reference standard: following index test

Aktan 2012 *(Continued)*

	Training of examiner: calibrated Blinding to index test: yes Multiple tests: no Site selection: sectioned teeth Target condition: healthy, enamel, dentinal		
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0		
Comparative			
Notes			
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			

Aktan 2012 *(Continued)*

DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Almosa 2014

Study characteristics	
Patient Sampling	Method of sampling: consecutive Included conditions: no cavitation and enamel lesions Teeth: permanent premolars and anterior - buccal Sealants: unclear Surface: smooth

Almosa 2014 *(Continued)*

Patient characteristics and setting	Age: mean 22.5 years Sex: 33 male, 56 female Ethnicity: not reported Country: Saudi Arabia Setting: governmental and private orthodontic clinics Number of participants/teeth/sites: 89/822/1653 Prevalence: enamel 0.33, dentine 0.01
Index tests	Category of test: DIAGNOdent pen Sequence of test(s): visual and DIAGNOdent pen conducted consecutively Examiner training and calibration: training workshop attended Teeth cleaning prior to examination: cleaned with rubber cup, pumice paste, and floss Tooth drying prior to examination: dried with compressed air Threshold applied: 0-13 sound, 14-20 enamel (outer), 21-29 enamel (deep), > 30 dentinal Device specifics: flat tip
Target condition and reference standard(s)	Category: visual (ICDAS) Sequence of index test and reference standard: consecutively with index test Training of examiner: training workshop Blinding to index test: no Multiple tests: no Site selection: unclear Target condition: sound = ICDAS 0, enamel = ICDAS 1 and 2, deep enamel = ICDAS 3 and 4, dentine = ICDAS 5 and 6
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0
Comparative	
Notes	Reference standard classify ICDAS 3 and 4 as enamel caries which conflicts other definitions
Methodological quality	

Almosa 2014 *(Continued)*

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	No		

Almosa 2014 *(Continued)*

Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			High
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Alomari 2015

Study characteristics	
Patient Sampling	Method of sampling: selected Included conditions: no cavitation and enamel lesions Teeth: permanent premolars and molars Sealants: unclear Surface: occlusal
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported Country: Kuwait Setting: extracted teeth Number of participants/teeth/sites: 160 teeth Prevalence: enamel 0.89, dentine 0.38
Index tests	Category of test: combined visual, radiograph, and DIAGNOdent Sequence of test(s): examination 1: visual only, examination 2: visual with radiographs, examination 3: visual, radiographs, and DIAGNOdent Examiner training and calibration: yes Teeth cleaning prior to examination: prophylaxis brush using pumice slurry

Alomari 2015 *(Continued)*

	Tooth drying prior to examination: dried for 5 seconds Threshold applied: manufacturer's instructions Device specifics: tip not reported
Target condition and reference standard(s)	Category: histology Sequence of index test and reference standard: following index test Training of examiner: calibration performed Blinding to index test: unclear Multiple tests: no Site selection: highest score from sectioned tooth Target condition: (Downer): sound, outer half of the enamel, inner half of the enamel, outer half of the dentine, inner half of the dentine
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: 1 month Participants receiving both tests but excluded from results: 0
Comparative	
Notes	Data for the enamel caries threshold provided by author

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		

Alomari 2015 *(Continued)*

If multiple tests were applied were different examiners used for each (in vivo)?			
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			High
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?			
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			High
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Alwas-Danowska 2002

Study characteristics	
Patient Sampling	Method of sampling: selected - participants volunteered Included conditions: no cavitation and enamel lesions Teeth: permanent premolars and molars Sealants: unclear Surface: occlusal
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported Country: Poland Setting: extracted teeth Number of participants/teeth/sites: 50 teeth Prevalence: unclear
Index tests	Category of test: DIAGNOdent completed in vivo and vitro, but no reference standard on the in vivo assessment Sequence of test(s): index test before reference standard Examiner training and calibration: yes Teeth cleaning prior to examination: not reported Tooth drying prior to examination: not reported Threshold applied: "The cut-off for the DIAGNOdent was between values 20 and 21" Device specifics: tip A
Target condition and reference standard(s)	Category: histology Sequence of index test and reference standard: following index test Training of examiner: not reported Blinding to index test: yes Multiple tests: no Site selection: sectioned teeth Target condition: lesion depth in mm, unlear threshold
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: unclear

Alwas-Danowska 2002 *(Continued)*

	Participants receiving both tests but excluded from results: 0		
Comparative			
Notes	Cannot extract data for 2x2 table as prevalence is not reported		
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Unclear		
If multiple tests were applied were different examiners used for each (in vivo)?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		Unclear risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Unclear		
If multiple tests were applied were different examiners used for each (in vivo)?	Unclear		

Alwas-Danowska 2002 *(Continued)*

Could the conduct or interpretation of the index test have introduced bias?		Unclear risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	No		
Could the patient flow have introduced bias?		Unclear risk	

Angnes 2005

Study characteristics	
Patient Sampling	Method of sampling: selected Included conditions: no cavitation and enamel lesions Teeth: permanent third molars Sealants: no Surface: occlusal
Patient characteristics and setting	Age: 19 to 35 years Sex: not reported Ethnicity: not reported Country: Brazil Setting: adult volunteers - "38 adult volunteers (19–35 years old) from Joaçaba, SC, Brazil, who had at least one third molar indicated for extraction" Number of participants/teeth/sites: 38/57/110

Angnes 2005 *(Continued)*

	Prevalence: 0.82 enamel, 0.14 dentine
Index tests	Category of test: DIAGNOdent Sequence of test(s): DIAGNOdent, visual, and radiography before reference standard Examiner training and calibration: yes, "Two examiners participated in this study; one of them trained the other on diagnostic procedures" Teeth cleaning prior to examination: rotating bristle brush Tooth drying prior to examination: 3-second air spray Threshold applied: < 15 sounds and early enamel, 15-19 late enamel and early dentine, > 19 deep dentine, analysis performed at > 19 level Device specifics: tip not reported
Target condition and reference standard(s)	Category: combined test of: visual, drill, radiograph Sequence of index test and reference standard: visual element completed before DIAGNOdent Training of examiner: not reported Blinding to index test: no Multiple tests: yes Site selection: not clearly reported Target condition: sound, inactive enamel, active enamel, dentinal
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0
Comparative	
Notes	Used data from first examiner

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		

Angnes 2005 *(Continued)*

Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			

Angnes 2005 *(Continued)*

Was there an appropriate interval between index test and reference standard?	Yes	
Did all patients receive the same reference standard?	Yes	
Were all patients included in the analysis?	Yes	
Could the patient flow have introduced bias?		Low risk

Anttonen 2003

Study characteristics	
Patient Sampling	Method of sampling: consecutive Included conditions: no cavitation and enamel lesions Teeth: primary molars and premolars Sealants: yes Surface: occlusal
Patient characteristics and setting	Age: 7 to 8 years Sex: not reported Ethnicity: not reported Country: Finland Setting: public dental clinics Number of participants/teeth/sites: 55 participants/650 teeth Prevalence: not reported
Index tests	Category of test: DIAGNOdent Sequence of test(s): visual then DIAGNOdent, then drilling and radiographs Examiner training and calibration:yes Teeth cleaning prior to examination: not reported Tooth drying prior to examination: air syringe Threshold applied: at intervals of 10 from 0-100 Device specifics: tip not reported
Target condition and reference standard(s)	Category: combined test of: visual, drill, radiograph Sequence of index test and reference standard: visual element completed before DIAGNOdent Training of examiner: not reported Blinding to index test: no

Anttonen 2003 *(Continued)*

	Multiple tests: yes Site selection: not clearly reported Target condition: sound, inactive enamel, active enamel, dentinal
Flow and timing	Participants with index test but no reference standard: unclear Participants with reference standard but no index test: unclear Time interval between tests: minimal Participants receiving both tests but excluded from results: unclear
Comparative	
Notes	Unclear reporting of data. Primary teeth had visual and DIAGNOdent only. Permanent had excavation and radiograph, but unclear on numbers of who receiving tests

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	No		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	No		
If a threshold was used, was it pre-specified?	No		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			High
DOMAIN 2: Index Test (Green fluorescence)			

Anttonen 2003 *(Continued)*

DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	No		
If a threshold was used, was it pre-specified?	No		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			High
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	No		
Were the reference standard results interpreted without knowledge of the results of the index tests?	No		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			High
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	No		
Were all patients included in the analysis?	Unclear		
Could the patient flow have introduced bias?		High risk	

Apostolopoulou 2009

Study characteristics	
Patient Sampling	Method of sampling: selected Included conditions: no cavitation and early lesions Teeth: primary molars Sealants: no Surface: occlusal

Apostolopoulou 2009 *(Continued)*

Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported Country: Greece Setting: extracted teeth Number of participants/teeth/sites: 24 teeth/111 sites Prevalence: enamel 0.98, dentine 0.22
Index tests	Category of test: DIAGNOdent Sequence of test(s): index tests (visual, radiograph, and DIAGNOdent) performed prior to reference standard Examiner training and calibration: not reported Teeth cleaning prior to examination: toothbrush and pumice-free paste Tooth drying prior to examination: air dried 5 seconds Threshold applied: converted scale, unclear, "the original DD readings on the 0-99 scale were converted, using Cronbach's A coefficient alpha, to the 0, 1 and 2 caries scoring scale used by all other methods" Device specifics: tip A
Target condition and reference standard(s)	Category: histology Sequence of index test and reference standard: index test then reference standard Training of examiner: not reported Blinding to index test: unclear Multiple tests: no Site selection: sectioned teeth Target condition: sound, enamel, dentine
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0
Comparative	
Notes	
Methodological quality	

Apostolopoulou 2009 *(Continued)*

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Unclear		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		Unclear risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Unclear		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		Unclear risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		

Apostolopoulou 2009 *(Continued)*

Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Arslan 2014

Study characteristics	
Patient Sampling	Method of sampling: selected Included conditions: unclear "suspected to have occlusal caries lesions" Teeth: permanent premolars and molars Sealants: not reported Surface: occlusal
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported Country: Turkey Setting: extracted teeth Number of participants/teeth/sites: 60 teeth Prevalence: enamel 0.82, dentine 0.45
Index tests	Category of test: DIAGNOdent pen Sequence of test(s): index tests (visual, radiograph, DIAGNOdent pen, micro-computed tomography examination) performed prior to reference standard Examiner training and calibration: 2 experienced examiners Teeth cleaning prior to examination: not reported

Arslan 2014 *(Continued)*

	Tooth drying prior to examination: air dried 5 seconds Threshold applied: not reported Device specifics: not reported
Target condition and reference standard(s)	Category: histology Sequence of index test and reference standard: index test then reference standard Training of examiner: experienced examiners Blinding to index test: same examiners as index test Multiple tests: no Site selection: sectioned teeth Target condition: sound, enamel, dentine
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Unclear		

Arslan 2014 *(Continued)*

If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		Unclear risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Unclear
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Unclear		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		Unclear risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Unclear
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	No		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Attrill 2001

Study characteristics	
Patient Sampling	Method of sampling: selected Included conditions: no cavitation and early lesions Teeth: primary molars Sealants: no Surface: occlusal
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported Country: UK Setting: extracted teeth Number of participants/teeth/sites: 58 teeth Prevalence: enamel 0.60, dentine 0.52
Index tests	Category of test: DIAGNOdent Sequence of test(s): index tests (visual, DIAGNOdent, and radiograph) performed prior to reference standard Examiner training and calibration: none, experienced examiners Teeth cleaning prior to examination: "cleaned with a pumice and water slurry" Tooth drying prior to examination: not reported Threshold applied: 0–9 sound/early enamel caries, 10–17 enamel caries, 18–99 dentinal caries Device specifics: tip A
Target condition and reference standard(s)	Category: histology Sequence of index test and reference standard: index test then reference standard Training of examiner: not reported Blinding to index test: yes Multiple tests: no Site selection: sectioned teeth Target condition: sound, enamel, dentine (outer third), dentine (mid and inner)
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: minimal

Attrill 2001 *(Continued)*

	Participants receiving both tests but excluded from results: 0		
Comparative			
Notes	A threshold was applied to the index test which categorised early enamel caries with sound surfaces, therefore the data are not appropriate for meta-analysis		
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	No		

Attrill 2001 *(Continued)*

Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Bahrololoomi 2015

Study characteristics	
Patient Sampling	Method of sampling: selected Included conditions: no cavitation and early lesions "intact or had incipient and inconspicuous caries with or without colour change were selected" Teeth: permanent molars Sealants: no Surface: occlusal
Patient characteristics and setting	Age: 7 to 13 years Sex: not reported Ethnicity: not reported Country: Iran Setting: dental school

Bahrololoomi 2015 *(Continued)*

	Number of participants/teeth/sites: 31 participants/115 teeth (6 of these were excluded "due to patient dropout" so they became 109 teeth) Prevalence: enamel 0.94, dentine 0.37
Index tests	Category of test: DIAGNOdent Sequence of test(s): index tests (visual, radiograph, DIAGNOdent) performed prior to reference standard Examiner training and calibration: not reported Teeth cleaning prior to examination: cleaning with a rubber cup and pumice powder Tooth drying prior to examination: isolation with cotton rolls, and drying Threshold applied: defined in study: 0-7 sound, 8-10 enamel, 11+ dentine Device specifics: not reported
Target condition and reference standard(s)	Category: excavation - in cases with obvious or ambiguous caries Sequence of index test and reference standard: index test then reference standard Training of examiner: not reported Blinding to index test: unclear Multiple tests: no Site selection: excavated suspicious site Target condition: sound, enamel, dentine
Flow and timing	Participants with index test but no reference standard: unclear whether all surfaces were excavated and if not then what the reference standard was Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0
Comparative	
Notes	Examiner 2 results used for analysis

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		

Bahrololoomi 2015 *(Continued)*

Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Unclear
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Unclear
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Unclear risk	

Bahrololoomi 2015 *(Continued)*

Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Unclear		
Were all patients included in the analysis?	No		
Could the patient flow have introduced bias?		High risk	

Bamzahim 2002

Study characteristics	
Patient Sampling	Method of sampling: selected Included conditions: no cavitation and early lesions Teeth: permanent premolars Sealants: no Surface: occlusal
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported Country: Sweden Setting: extracted teeth from orthodontic patients Number of participants/teeth/sites: 87 teeth Prevalence: enamel 0.78, dentine 0.26
Index tests	Category of test: DIAGNOdent Sequence of test(s): index tests (DIAGNOdent then ECM) performed prior to reference standard Examiner training and calibration: not reported Teeth cleaning prior to examination: cleaned with toothbrush and scaled Tooth drying prior to examination: air dried for 10 seconds Threshold applied: 18+ dentine; other thresholds not reported Device specifics: conical tip
Target condition and reference standard(s)	Category: histology

Bamzahim 2002 *(Continued)*

	Sequence of index test and reference standard: index test then reference standard Training of examiner: not reported Blinding to index test: yes Multiple tests: no Site selection: sectioned teeth and location marked on photograph Target condition: sound, enamel, dentine (outer third), dentine (mid and inner)
Flow and timing	Participants with index test but no reference standard: 10 Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0
Comparative	
Notes	Data not available at enamel level

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Unclear		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		Unclear risk	

Bamzahim 2002 *(Continued)*

Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Unclear		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		Unclear risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	No		
Could the patient flow have introduced bias?		High risk	

Bamzahim 2004

Study characteristics	
Patient Sampling	Method of sampling: selected

Bamzahim 2004 *(Continued)*

	Included conditions: no cavitation and early lesions "suspicious sites", restoration also included Teeth: permanent premolars and molars Sealants: no Surface: unclear, study investigating secondary caries
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported Country: Sweden Setting: extracted teeth Number of participants/teeth/sites: 87 teeth Prevalence: enamel 0.52
Index tests	Category of test: DIAGNOdent Sequence of test(s): index tests (DIAGNOdent then radiograph) performed prior to reference standard Examiner training and calibration: not reported Teeth cleaning prior to examination: wiped with paper towel Tooth drying prior to examination: air dried for 10 seconds Threshold applied: on threshold of 20 was applied for generating sensitivity and specificity, ROC curves were generated according to thresholds: 1 = values ranging from 0 to 10, 2 = values ranging from 11 to 20, 3 = values ranging from 21 to 30, 4 = values ranging from 31 to 40, 5 = values above 40 Device specifics: conical tip
Target condition and reference standard(s)	Category: excavation of restorative material followed by histology Sequence of index test and reference standard: index test then reference standard Training of examiner: not reported Blinding to index test: yes Multiple tests: no Site selection: sectioned teeth and location marked on photograph Target condition: soft or hard
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: minimal

Bamzahim 2004 *(Continued)*

	Participants receiving both tests but excluded from results: 0		
Comparative			
Notes			
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	

Bamzahim 2004 *(Continued)*

Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Barberia 2008

Study characteristics	
Patient Sampling	Method of sampling: consecutive Included conditions: unclear Teeth: primary and permanent molars Sealants: no Surface: occlusal
Patient characteristics and setting	Age: 6 to 14 years Sex: not reported Ethnicity: not reported Country: Spain Setting: attending dental clinic Number of participants/teeth/sites: 320 teeth Prevalence: enamel 0.22, dentine 0.08
Index tests	Category of test: DIAGNOdent

Barberia 2008 *(Continued)*

	Sequence of test(s): index tests performed after reference standard, but examiners were blind to visual examination Examiner training and calibration: not reported Teeth cleaning prior to examination: not reported Tooth drying prior to examination: 2 seconds Threshold applied: 0-4 healthy, 5-25 enamel, 26+ dentine Device specifics: "same tip used for all"
Target condition and reference standard(s)	Category: visual Sequence of index test and reference standard: index test then reference standard Training of examiner: no but experienced examiner Blinding to index test: yes Multiple tests: no Site selection: unclear Target condition: no treatment required, potential for remineralisation, restoration required
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (All)			

Barberia 2008 *(Continued)*

Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?			
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?			
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	No		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		

Barberia 2008 *(Continued)*

Could the patient flow have introduced bias?	Low risk

Baseren 2003

Study characteristics	
Patient Sampling	Method of sampling: selected Included conditions: no cavitation and early lesions Teeth: permanent molars Sealants: not reported Surface: occlusal
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported Country: Turkey Setting: extracted teeth Number of participants/teeth/sites: 35 teeth Prevalence: enamel 0.39, dentine 0.19
Index tests	Category of test: DIAGNOdent Sequence of test(s): index tests (DIAGNOdent) performed prior to reference standard Examiner training and calibration: calibrated examiners Teeth cleaning prior to examination: water, brush, and pumice Tooth drying prior to examination: paper tissues Threshold applied: 0-13 sound, 14-19 enamel, > 20 dentine Device specifics: tip A
Target condition and reference standard(s)	Category: histology Sequence of index test and reference standard: index test then reference standard Training of examiner: not reported Blinding to index test: unclear Multiple tests: no Site selection: location marked on drawing

Baseren 2003 *(Continued)*

	Target condition: sound, enamel, dentine
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?			
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			

Baseren 2003 *(Continued)*

Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?			
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Bengtson 2005

Study characteristics	
Patient Sampling	Method of sampling: selected Included conditions: no cavitation and early lesions Teeth: primary molars Sealants: not reported Surface: occlusal
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported

Bengtson 2005 *(Continued)*

	Country: Brazil Setting: extracted teeth Number of participants/teeth/sites: 50 teeth/87 surfaces Prevalence: enamel 0.53, dentine 0.06
Index tests	Category of test: DIAGNOdent Sequence of test(s): index tests (visual then DIAGNOdent) performed prior to reference standard Examiner training and calibration: no training Teeth cleaning prior to examination: water/pumice slurry Tooth drying prior to examination: compressed air for 10 seconds Threshold applied: 0-4 sound, 5-12 enamel, > 12 dentine Device specifics: tip A
Target condition and reference standard(s)	Category: histology Sequence of index test and reference standard: following index test Training of examiner: not reported Blinding to index test: yes Multiple tests: no Site selection: marked on a drawing Target condition: sound, initial enamel, advanced enamel, initial dentine, advanced dentine
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		

Bengtson 2005 *(Continued)*

Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern

Bengtson 2005 *(Continued)*

DOMAIN 4: Flow and Timing		
Was there an appropriate interval between index test and reference standard?	Yes	
Did all patients receive the same reference standard?	Yes	
Were all patients included in the analysis?	Yes	
Could the patient flow have introduced bias?		Low risk

Bittar 2012

Study characteristics	
Patient Sampling	Method of sampling: selected Included conditions: no cavitation and early lesions Teeth: primary molars Sealants: not reported Surface: occlusal and approximal
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported Country: Brazil Setting: extracted teeth Number of participants/teeth/sites: 33 teeth/55 surfaces Prevalence: occlusal - enamel 0.67, dentine 0.22 approximal - enamel 0.4, dentine 0.28
Index tests	Category of test: DIAGNOdent pen, "authors tried to reproduce the contact points as best as possible, placing the teeth in arch models" Sequence of test(s): index tests (DIAGNOdent pen) performed prior to reference standard Examiner training and calibration: experienced, calibrated dentist Teeth cleaning prior to examination: brush and slurry Tooth drying prior to examination: not clearly reported Threshold applied: 0-8 sound, 9-30 enamel, > 31 dentine Device specifics: tip 2

Bittar 2012 *(Continued)*

Target condition and reference standard(s)	Category: histology Sequence of index test and reference standard: following index test Training of examiner: 2 experienced examiners Blinding to index test: yes Multiple tests: no Site selection: marked on a drawing Target condition: sound, initial enamel, advanced enamel, intial dentine, advanced dentine
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	

Bittar 2012 *(Continued)*

Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Bizhang 2016

Study characteristics	
Patient Sampling	Method of sampling: consecutive Included conditions: no cavitation and early lesions

Bizhang 2016 *(Continued)*

	Teeth: permanent molars Sealants: not reported Surface: approximal
Patient characteristics and setting	Age: 18 to 65 years, mean 26.7 Sex: not reported Ethnicity: not reported Country: Germany Setting: in vivo with recruited patients but setting unclear Number of participants/teeth/sites: 20 teeth/341 surfaces Prevalence: not reported
Index tests	Category of test: DIAGNOdent pen Sequence of test(s): index tests (visual, radiograph, then DIAGNOdent pen) where radiograph is the reference standard Examiner training and calibration: calibrated dentist Teeth cleaning prior to examination: oral prophylaxis and floss Tooth drying prior to examination: 5 seconds compressed air Threshold applied: > 16 dentine Device specifics: not reported
Target condition and reference standard(s)	Category: radiograph Sequence of index test and reference standard: radiographs performed at screening session, then again 1 week later Training of examiner: not reported Blinding to index test: unclear Multiple tests: no Site selection: all approximal surfaces Target condition: sound, initial enamel, advanced enamel, intial dentine, advanced dentine
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: one week Participants receiving both tests but excluded from results: 0

Bizhang 2016 *(Continued)*

Comparative
Notes

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			Unclear
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern

Bizhang 2016 *(Continued)*

DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	No		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Boston 2003

Study characteristics	
Patient Sampling	Method of sampling: selected Included conditions: no cavitation, early lesions, and restorations Teeth: permanent incisors, canines, premolar, and molar Sealants: not reported Surface: occlusal
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported Country: US Setting: extracted teeth Number of participants/teeth/sites: 15 teeth/30 surfaces Prevalence: enamel 0.57, dentine 0.37
Index tests	Category of test: DIAGNOdent Sequence of test(s): index tests (visual and DIAGNOdent) performed prior to reference standard

Boston 2003 *(Continued)*

	Examiner training and calibration: not reported Teeth cleaning prior to examination: water and slurry Tooth drying prior to examination: "air blast for 10 seconds" Threshold applied: "calculated 'best' threshold" Device specifics: tip A
Target condition and reference standard(s)	Category: histology Sequence of index test and reference standard: index test then reference standard Training of examiner: not reported Blinding to index test: yes Multiple tests: no Site selection: marked on model of tooth Target condition: sound, enamel, dentine
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0
Comparative	
Notes	Primary outcome is secondary caries

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		

Boston 2003 *(Continued)*

If a threshold was used, was it pre-specified?	No		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Bozdemir 2013

Study characteristics	
Patient Sampling	Method of sampling: unclear Included conditions: no cavitation and early lesions Teeth: permanent molars Sealants: no Surface: occlusal
Patient characteristics and setting	Age: 20 to 25 years, mean 20.2 Sex: 7 male, 30 female Ethnicity: not reported Country: Turkey Setting: dental school Number of participants/teeth/sites: 37 teeth/156 surfaces Prevalence: not reported, only those suspected of caries received reference standard so data do not reflect full sample
Index tests	Category of test: DIAGNOdent pen and Midwest Sequence of test(s): index tests performed prior to reference standard, same 2 examiners completed all tests Examiner training and calibration: "previously trained" Teeth cleaning prior to examination: rubber cup and paste Tooth drying prior to examination: air dried for 3 seconds Threshold applied: DIAGNOdent pen: 0–13 healthy, 14–20 enamel, 21–29 superficial dentine, > 30 deep dentine Midwest: no signal/green light - healthy, slow signal/red light - enamel, medium signal/red light - superficial dentine, rapid or continuous signal/red light - deep dentine Device specifics: DIAGNOdent pen - cone shaped tip
Target condition and reference standard(s)	Category: excavation - only those diagnosed as having caries by index tests were investigated Sequence of index test and reference standard: index test then reference standard Training of examiner: not reported Blinding to index test: yes Multiple tests: no Site selection: complete occlusal fissure

Bozdemir 2013 *(Continued)*

	Target condition: sound, enamel (outer), enamel (inner), dentine (outer), dentine (inner)
Flow and timing	Participants with index test but no reference standard: 30 Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0
Comparative	
Notes	Data not used because 156 sites were included but only 126 were opened

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		

Bozdemir 2013 *(Continued)*

If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	No		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	No		
Were all patients included in the analysis?	No		
Could the patient flow have introduced bias?		High risk	

Braga 2006

Study characteristics	
Patient Sampling	Method of sampling: selected Included conditions: no cavitation and early lesions Teeth: primary molars Sealants: no Surface: occlusal
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported Country: Brazil

Braga 2006 *(Continued)*

	Setting: extracted teeth Number of participants/teeth/sites: 54 teeth Prevalence: not reported
Index tests	Category of test: DIAGNOdent Sequence of test(s): index tests performed prior to reference standard Examiner training and calibration: not reported Teeth cleaning prior to examination: brush and slurry Tooth drying prior to examination: air dried for 3 seconds Threshold applied: 0-9 sound, 10-17 enamel, 18-99 dentine Device specifics: tip A
Target condition and reference standard(s)	Category: histology Sequence of index test and reference standard: index test then reference standard Training of examiner: not reported Blinding to index test: unclear Multiple tests: no Site selection: sectioned teeth Target condition: sound, enamel (outer), enamel (inner), dentine (outer), dentine (inner)
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		

Braga 2006 *(Continued)*

Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?			
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?			
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern

Braga 2006 *(Continued)*

DOMAIN 4: Flow and Timing		
Was there an appropriate interval between index test and reference standard?	Yes	
Did all patients receive the same reference standard?	Yes	
Were all patients included in the analysis?	Yes	
Could the patient flow have introduced bias?		Low risk

Braga 2007

Study characteristics	
Patient Sampling	Method of sampling: selected Included conditions: no cavitation and early lesions Teeth: permanent molars Sealants: not reported Surface: occlusal
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported Country: Brazil Setting: extracted teeth Number of participants/teeth/sites: 86 teeth/123 surfaces Prevalence: not reported
Index tests	Category of test: DIAGNOdent Sequence of test(s): index tests performed prior to reference standard Examiner training and calibration: not reported Teeth cleaning prior to examination: brush and slurry Tooth drying prior to examination: air dried for 3 seconds Threshold applied: compared multiple thresholds Device specifics: tip A
Target condition and reference standard(s)	Category: histology Sequence of index test and reference standard: index test then reference standard

Braga 2007 *(Continued)*

	Training of examiner: not reported Blinding to index test: unclear Multiple tests: no Site selection: sectioned teeth Target condition: sound, enamel (outer), enamel (inner), dentine (outer), dentine (inner)
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern

Braga 2007 *(Continued)*

DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Braga 2008

Study characteristics	
Patient Sampling	Method of sampling: selected Included conditions: no cavitation and early lesions Teeth: primary molars Sealants: no

Braga 2008 *(Continued)*

	Surface: occlusal
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported Country: Brazil Setting: extracted teeth Number of participants/teeth/sites: 132 teeth/181 sites Prevalence: not reported
Index tests	Category of test: DIAGNOdent Sequence of test(s): index tests performed prior to reference standard Examiner training and calibration: not reported Teeth cleaning prior to examination: brush and slurry Tooth drying prior to examination: air dried for 3 seconds Threshold applied: calculated within study, 0-5 sound, 6-10 outer enamel, 11-15 inner enamel, 16+ dentine Device specifics: tip A
Target condition and reference standard(s)	Category: histology Sequence of index test and reference standard: index test then reference standard Training of examiner: not reported Blinding to index test: unclear Multiple tests: no Site selection: sectioned teeth Target condition: sound, enamel (outer), enamel (inner), dentine (outer), dentine (inner)
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0
Comparative	
Notes	
Methodological quality	

Braga 2008 *(Continued)*

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		
If multiple tests were applied were different examiners used for each (in vivo)?			
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		
If multiple tests were applied were different examiners used for each (in vivo)?			
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		

Braga 2008 *(Continued)*

Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Braga 2009

Study characteristics	
Patient Sampling	Method of sampling: selected Included conditions: no cavitation and early lesions Teeth: primary molars Sealants: no Surface: occlusal
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported Country: Brazil Setting: extracted teeth Number of participants/teeth/sites: 84 participants/131 sites Prevalence: enamel 0.63, dentine 0.26
Index tests	Category of test: DIAGNOdent pen Sequence of test(s): index tests (visual, radiograph, laser fluorescence) prior to reference standard Examiner training and calibration: yes, trained Teeth cleaning prior to examination: brush and slurry Tooth drying prior to examination: air dried for 3 seconds

Braga 2009 *(Continued)*

	Threshold applied: calculated within study, 0-4 sound, 4.1-38 white spot, 38+ cavitated Device specifics: tip 1
Target condition and reference standard(s)	Category: histology Sequence of index test and reference standard: index test then reference standard Training of examiner: not reported Blinding to index test: unclear Multiple tests: no Site selection: sectioned teeth Target condition: sound, enamel (outer), enamel (inner), dentine (outer), dentine (inner)
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		

Braga 2009 *(Continued)*

If multiple tests were applied were different examiners used for each (in vivo)?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		
If multiple tests were applied were different examiners used for each (in vivo)?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Burin 2005

Study characteristics	
Patient Sampling	Method of sampling: selected Included conditions: no cavitation and early lesions Teeth: permanent premolars and molars Sealants: no Surface: occlusal
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported Country: Brazil Setting: extracted teeth Number of participants/teeth/sites: 54 participants/105 sites Prevalence: enamel 0.71, dentine 0.17
Index tests	Category of test: DIAGNOdent Sequence of test(s): index tests performed prior to reference standard Examiner training and calibration: not reported Teeth cleaning prior to examination: unclear Tooth drying prior to examination: unclear Threshold applied: calculated within study, 0-11 sound and outer enamel, 12-16 inner enamel, 16+ dentine Device specifics: tip A
Target condition and reference standard(s)	Category: histology Sequence of index test and reference standard: index test then reference standard Training of examiner: not reported Blinding to index test: unclear Multiple tests: no Site selection: sectioned teeth Target condition: healthy, enamel, up to EDJ, dentine
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: minimal

Burin 2005 *(Continued)*

	Participants receiving both tests but excluded from results: 0		
Comparative			
Notes	Data only available at dentine level		
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		
If multiple tests were applied were different examiners used for each (in vivo)?			
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		
If multiple tests were applied were different examiners used for each (in vivo)?			
Could the conduct or interpretation of the index test have introduced bias?		High risk	

Burin 2005 *(Continued)*

Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Bussaneli 2015

Study characteristics	
Patient Sampling	Method of sampling: selected Included conditions: no cavitation and early lesions Teeth: permanent premolars and molars Sealants: no Surface: occlusal
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported Country: Brazil Setting: extracted teeth Number of participants/teeth/sites: 102 teeth Prevalence: enamel 0.70, dentine 0.19
Index tests	Category of test: DIAGNOdent pen and QLF Inspektor

Bussaneli 2015 *(Continued)*

	Sequence of test(s): index tests (radiograph, near infrared then DIAGNOdent pen and QLF) prior to reference standard Examiner training and calibration: experienced Teeth cleaning prior to examination: unclear Tooth drying prior to examination: unclear Threshold applied: DIAGNOdent pen: 0-14, 15-21 enamel, 22-37 outer dentine, 38+ deep dentine QLF Inspektor: ΔF values were characterized as follows: −0.5 to −10, healthy; −10.5 to −35, enamel carious lesions; and −35.5 to −45, cavitated lesion with visible dentine Device specifics: DIAGNOdent pen: cylindrical sapphire QLF Inspektor: analysed using Inspektor™ Pro software (version 2.0.0.32)
Target condition and reference standard(s)	Category: histology Sequence of index test and reference standard: index test then reference standard Training of examiner: not reported Blinding to index test: unclear Multiple tests: no Site selection: sectioned teeth Target condition: healthy, enamel, lesion at the dentino-enamel junction or dentinal
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 8 teeth excluded from results as near infrared device failed to return a result, therefore excluded from all tests
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		

Bussaneli 2015 *(Continued)*

Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	

Bussaneli 2015 *(Continued)*

Are there concerns that the target condition as defined by the reference standard does not match the question?		Low concern
DOMAIN 4: Flow and Timing		
Was there an appropriate interval between index test and reference standard?	Yes	
Did all patients receive the same reference standard?	Yes	
Were all patients included in the analysis?	No	
Could the patient flow have introduced bias?		High risk

Bussaneli 2015a

Study characteristics	
Patient Sampling	Method of sampling: selected Included conditions: no cavitation and early lesions "sound or carious primary molars in proximal contact", exclusions "Teeth with restoration, occlusal caries, hypoplasias, and an advanced stage of rhizolysis were not included" Teeth: primary molars Sealants: not reported Surface: approximal
Patient characteristics and setting	Age: 5 to 9 years Sex: not reported Ethnicity: not reported Country: Brazil Setting: extracted teeth Number of participants/teeth/sites: 45 participants/59 teeth Prevalence: enamel 0.71, dentine 0.58
Index tests	Category of test: DIAGNOdent pen Sequence of test(s): index tests (visual, DIAGNOdent pen, radiograph) prior to reference standard Examiner training and calibration: experienced Teeth cleaning prior to examination: brush at low speed, using prophylactic paste and dental floss Tooth drying prior to examination: unclear Threshold applied: 0-14 sound, 15-21 enamel, 22-37 outer dentine, 38+ inner dentine

Bussaneli 2015a *(Continued)*

	Device specifics: tip 1
Target condition and reference standard(s)	Category: visual after separation using orthodontic rubber bands (4 mm) for 7 days Sequence of index test and reference standard: index test then reference standard Training of examiner: 2 trained and experienced examiners Blinding to index test: unclear Multiple tests: no Site selection: all approximal surfaces Target condition: healthy, active lesions without loss of structure, signs of caries requiring restoration
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	No		

Bussaneli 2015a *(Continued)*

Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	No		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			High
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Castilho 2016

Study characteristics

Castilho 2016 *(Continued)*

Patient Sampling	Method of sampling: consecutive Included conditions: no cavitation and early lesions Teeth: third molars, requiring extraction Sealants: no Surface: occlusal
Patient characteristics and setting	Age: 16 to 39 years Sex: 10 male, 16 female Ethnicity: not reported Country: Brazil Setting: dental clinic Number of participants/teeth/sites: 26 participants/43 teeth Prevalence: enamel 0.81, dentine 0.07
Index tests	Category of test: DIAGNOdent Sequence of test(s): index tests (visual then DIAGNOdent pen) prior to reference standard Examiner training and calibration: yes Teeth cleaning prior to examination: brush and slurry Tooth drying prior to examination: yes Threshold applied: 0-5 sound, 6-14 outer enamel, 15-20 inner enamel, 21-99 dentine Device specifics: tip A
Target condition and reference standard(s)	Category: histology Sequence of index test and reference standard: index test then reference standard Training of examiner: not reported Blinding to index test: yes Multiple tests: no Site selection: sectioned teeth Target condition: healthy, enamel, dentine
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: minimal

Castilho 2016 *(Continued)*

	Participants receiving both tests but excluded from results: 0		
Comparative			
Notes			
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	

Castilho 2016 *(Continued)*

Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Chawla 2012

Study characteristics	
Patient Sampling	Method of sampling: selected Included conditions: no cavitation and early lesions Teeth: primary molars Sealants: not reported Surface: approximal
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported Country: Australia Setting: extracted teeth Number of participants/teeth/sites: 135 sites Prevalence: enamel 0.61, dentine 0.24
Index tests	Category of test: DIAGNOdent and DIAGNOdent pen

Chawla 2012 *(Continued)*

	Sequence of test(s): index tests (visual, radiograph, then DIAGNOdent then DIAGNOdent pen) prior to reference standard Examiner training and calibration: training completed Teeth cleaning prior to examination: brush and slurry Tooth drying prior to examination: yes Threshold applied: DIAGNOdent: 0-4 sound, 5-7 outer enamel, 8-10 inner enamel, 11-12 outer dentine, 13+ inner dentine DIAGNOdent pen: 0-4 sound, 5-8 outer enamel, 9-11 inner enamel, 12-15 outer dentine, 16+ inner dentine Device specifics: DIAGNOdent: tip B DIAGNOdent pen: angled tip
Target condition and reference standard(s)	Category: histology Sequence of index test and reference standard: index test then reference standard Training of examiner: not reported Blinding to index test: yes Multiple tests: no Site selection: sectioned teeth Target condition: healthy, inner/outer enamel, inner/outer dentine
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0
Comparative	
Notes	Data cannot inform the production of 2x2 table, so not included in analysis

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		

Chawla 2012 *(Continued)*

Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern

Chawla 2012 *(Continued)*

DOMAIN 4: Flow and Timing	
Was there an appropriate interval between index test and reference standard?	Yes
Did all patients receive the same reference standard?	Yes
Were all patients included in the analysis?	Yes
Could the patient flow have introduced bias?	Low risk

Chen 2012

Study characteristics	
Patient Sampling	Method of sampling: unclear Included conditions: no cavitation and early lesions Teeth: primary molars Sealants: not reported Surface: approximal
Patient characteristics and setting	Age: 5 to 9 years Sex: not reported Ethnicity: not reported Country: China Setting: dental hospital Number of participants/teeth/sites: 96 participants/216 teeth/256 sites Prevalence: enamel 0.50, dentine 0.35
Index tests	Category of test: DIAGNOdent Sequence of test(s): index tests (visual, DIAGNOdent then radiograph) prior to reference standard Examiner training and calibration: not reported Teeth cleaning prior to examination: brush, paste, floss Tooth drying prior to examination: 5 seconds air spray Threshold applied: calculated in study, 0-7 sound, 8-16 enamel, 17+ dentine; "Cut-off limits of DIAGNOdent pen were determined in a way that enabled highest sum of specificity and sensitivity" Device specifics: not reported
Target condition and reference standard(s)	Category: excavation or "direct evaluation" "depending on the examination findings, invasive treatments were per-

Chen 2012 *(Continued)*

	formed on cavitated molars." Not clear how many of the included surfaces received excavation and restoration Sequence of index test and reference standard: index test then reference standard Training of examiner: not reported Blinding to index test: no Multiple tests: yes Site selection: unclear Target condition: caries: cavities and white spots		
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0		
Comparative			
Notes			

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	

Chen 2012 *(Continued)*

Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	No		
Were the reference standard results interpreted without knowledge of the results of the index tests?	No		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	No		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Chong 2003

Study characteristics	
Patient Sampling	Method of sampling: selected Included conditions: no cavitation and early lesions

Chong 2003 *(Continued)*

	Teeth: permanent premolars Sealants: no Surface: occlusal
Patient characteristics and setting	Age: 12 to 15 years Sex: not reported Ethnicity: not reported Country: Australia Setting: extracted teeth Number of participants/teeth/sites: 320 teeth Prevalence: enamel 0.50, dentine 0.35
Index tests	Category of test: DIAGNOdent Sequence of test(s): index tests (visual, radiograph then DIAGNOdent) prior to reference standard Examiner training and calibration: not reported Teeth cleaning prior to examination: brush and slurry Tooth drying prior to examination: 5 seconds air spray Threshold applied: < 5 sound, 5-25 enamel, 26-35 dentine, > 35 advanced dentine Device specifics: not reported
Target condition and reference standard(s)	Category: visual Sequence of index test and reference standard: reference standard then index test Training of examiner: not reported Blinding to index test: yes Multiple tests: yes Site selection: not reported Target condition: visual: C0 sound, C1 no opacity, C2 opacity and not sticky, C3 opacity and sticky, C4 frank cavitation
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0
Comparative	

Chong 2003 *(Continued)*

Notes	Data used for fluorescence versus visual as this is the most clinically relevant, no sites identified as sound by index test

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	No		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	No		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern

Chong 2003 *(Continued)*

DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	No		
Were the reference standard results interpreted without knowledge of the results of the index tests?	No		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Cinar 2013

Study characteristics	
Patient Sampling	Method of sampling: selected Included conditions: no cavitation and early lesions Teeth: primary molars Sealants: no Surface: occlusal
Patient characteristics and setting	Age: 9 to 11 years Sex: not reported Ethnicity: not reported Country: Turkey Setting: dental hospital Number of participants/teeth/sites: 44 sites Prevalence: 0.75 enamel, 0.20 dentine
Index tests	Category of test: DIAGNOdent and DIAGNOdent pen Sequence of test(s): index tests (visual, radiograph then DIAGNOdent and DIAGNOdent pen) prior to reference standard

Cinar 2013 *(Continued)*

	Examiner training and calibration: not reported Teeth cleaning prior to examination: cleaned with paste Tooth drying prior to examination: not reported Threshold applied: DIAGNOdent: "manufacturer recommended" 0-5 sound, 6-14 outer enamel, 15-20 inner enamel, 21-99 dentine DIAGNOdent pen: 0-13 sound, 14-20 outer enamel, 21-29 inner enamel, 30+ dentine Device specifics: tip A
Target condition and reference standard(s)	Category: histology Sequence of index test and reference standard: index test then reference standard Training of examiner: not reported Blinding to index test: unclear Multiple tests: no Site selection: sectioned teeth Target condition: sound, outer enamel, inner enamel, dentine
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern

Cinar 2013 *(Continued)*

DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		

Cinar 2013 *(Continued)*

Were all patients included in the analysis?	Yes	
Could the patient flow have introduced bias?		Low risk

Costa 2002

Study characteristics	
Patient Sampling	Method of sampling: selected Included conditions: no cavitation and early lesions Teeth: permanent molars and premolars Sealants: excluded Surface: occlusal
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported Country: Brazil Setting: extracted teeth Number of participants/teeth/sites: 50 teeth Prevalence: 0.65 enamel, 0.31 dentine
Index tests	Category of test: DIAGNOdent Sequence of test(s): index tests (visual, DIAGNOdent, and radiograph) followed by reference standard Examiner training and calibration: yes Teeth cleaning prior to examination: yes Tooth drying prior to examination: not reported Threshold applied: 0-5 sound, 6-20 enamel, 21+ dentinal Device specifics: not reported
Target condition and reference standard(s)	Category: histology Sequence of index test and reference standard: index test then reference standard Training of examiner: not reported Blinding to index test: unclear Multiple tests: no Site selection: sectioned teeth

Costa 2002 *(Continued)*

	Target condition: sound, enamel, dentine
Flow and timing	Participants with index test but no reference standard: 1 (damaged during sectioning) Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			

Costa 2002 *(Continued)*

Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Costa 2007

Study characteristics	
Patient Sampling	Method of sampling: selected Included conditions: no cavitation and early lesions Teeth: primary and permanent, molars and premolars Sealants: not reported Surface: occlusal
Patient characteristics and setting	Age: 7 to 13 years Sex: not reported Ethnicity: not reported

Costa 2007 *(Continued)*

	Country: Brazil Setting: dental hospital Number of participants/teeth/sites: 55 teeth/564 teeth Prevalence: not reported
Index tests	Category of test: DIAGNOdent Sequence of test(s): index tests (visual, DIAGNOdent, and radiograph) followed by reference standard Examiner training and calibration: calibrated Teeth cleaning prior to examination: cleaned before visual examination Tooth drying prior to examination: cotton wool Threshold applied: 0-20 sound, 21-30 enamel, 31+ dentine Device specifics: not reported
Target condition and reference standard(s)	Category: excavation, unclear how many received this index test and which relied on the visual examination results Sequence of index test and reference standard: index test then reference standard Training of examiner: not reported Blinding to index test: unclear Multiple tests: no Site selection: if excavation was completed the whole site was investigated Target condition: sound, enamel, dentine
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0
Comparative	
Notes	Not possible to extract full 2x2 table

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		

Costa 2007 *(Continued)*

Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	No		
Were the reference standard results interpreted without knowledge of the results of the index tests?	No		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	

Costa 2007 *(Continued)*

Are there concerns that the target condition as defined by the reference standard does not match the question?		Low concern

DOMAIN 4: Flow and Timing		
Was there an appropriate interval between index test and reference standard?	Yes	
Did all patients receive the same reference standard?	No	
Were all patients included in the analysis?	Yes	
Could the patient flow have introduced bias?		Low risk

Diniz 2009

Study characteristics	
Patient Sampling	Method of sampling: selected "from sound to different degrees of non-cavitated caries lesions" Included conditions: no cavitation and early lesions Teeth: permanent first molars Sealants: no Surface: occlusal
Patient characteristics and setting	Age: 7 to 12 years Sex: not reported Ethnicity: not reported Country: Brazil Setting: extracted teeth Number of participants/teeth/sites: 35 participants/130 teeth Prevalence: enamel 0.89, dentine 0.67
Index tests	Category of test: DIAGNOdent Sequence of test(s): index tests (visual then DIAGNOdent) followed by reference standard Examiner training and calibration: experienced and trained Teeth cleaning prior to examination: pumice slurry and water Tooth drying prior to examination: air dried Threshold applied: 0-14 sound, 15-21 enamel, 22+ dentine Device specifics: not reported Note: different examiners for visual, DIAGNOdent, and reference standard

Diniz 2009 *(Continued)*

Target condition and reference standard(s)	Category: radiograph and visual (third dentist), excavation where appropriate Sequence of index test and reference standard: index test then reference standard Training of examiner: experienced Blinding to index test: yes Multiple tests: combined test Site selection: teeth were drawn to aid examiners Target condition: sound, inner/outer enamel, inner/outer dentine
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0
Comparative	
Notes	Cannot extract data at the enamel threshold so no included in meta-analysis

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	Yes		

Diniz 2009 *(Continued)*

Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	No		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	No		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		High risk	

Diniz 2011

Study characteristics

Diniz 2011 *(Continued)*

Patient Sampling	Method of sampling: selected "sound to carious were selected from a pool of extracted teeth" so unclear the level of cavity included Included conditions: no cavitation and early lesions Teeth: permanent third molars Sealants: not reported Surface: occlusal
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported Country: Brazil Setting: extracted teeth Number of participants/teeth/sites: 55 teeth Prevalence: enamel 0.89, dentine 0.11
Index tests	Category of test: DIAGNOdent, DIAGNOdent pen and VistaProof Sequence of test(s): index tests (visual inspected but not assessed then fluorescence devices) followed by reference standard Examiner training and calibration: 2 experienced examiners Teeth cleaning prior to examination: prophylactic paste using a slow-rotating contra angle handpiece with a Robinson brush (group 2) Tooth drying prior to examination: dried for 3 seconds Threshold applied: calculated within study DIAGNOdent: 0–15 sound, 16–25 enamel, 25+ dentine DIAGNOdent pen: 0–10 sound, 11–34 enamel, 34+ dentine VistaProof: 0–1.1 sound, 1.2–1.7 enamel, 1.7+ dentine Device specifics: VistaProof - specific software (DBSWIN) that translates the rates of red and green fluorescence into numbers corresponding to lesion severity
Target condition and reference standard(s)	Category: histology Sequence of index test and reference standard: index test then reference standard Training of examiner: 2 trained examiners Blinding to index test: unclear Multiple tests: no Site selection: marked on photographs then sectioned teeth Target condition: sound, inner/outer enamel, inner/outer dentine
Flow and timing	Participants with index test but no reference standard: 0

Diniz 2011 *(Continued)*

	Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0		
Comparative			
Notes	Results used from stage (2) after professional prophylaxis (prophylactic paste) for 10 seconds, rinsing for 3 seconds and drying for 3 seconds		

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		
If multiple tests were applied were different examiners used for each (in vivo)?	No		

Diniz 2011 *(Continued)*

Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Red fluorescence)			
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Diniz 2012

Study characteristics	
Patient Sampling	Method of sampling: selected "88 patients who each had at least one posterior tooth scheduled for extraction" Included conditions: no cavitation and early lesions "ranged from having macroscopically intact occlusal surfaces to different stages of non-cavitated and cavitated carious lesions" Teeth: permanent molars and premolars Sealants: excluded Surface: occlusal
Patient characteristics and setting	Age: 18 to 35 years Sex: not reported Ethnicity: not reported Country: Brazil

Diniz 2012 *(Continued)*

	Setting: extracted teeth Number of participants/teeth/sites: 88 participants/105 teeth Prevalence: enamel 0.95, dentine 0.26
Index tests	Category of test: DIAGNOdent, DIAGNOdent pen, and VistaProof Sequence of test(s): index tests (visual, radiograph, DIAGNOdent, DIAGNOdent pen, and VistaProof) followed by reference standard Examiner training and calibration: 1 experienced examiner - masked from results of fluorescence device Teeth cleaning prior to examination: low-speed handpiece with a rotating brush and water Tooth drying prior to examination: unclear Threshold applied: calculated within study DIAGNOdent: 0–15 sound, 16–25 enamel, 25+ dentine DIAGNOdent pen: 0–10 sound, 11–34 enamel, 34+ dentine VistaProof: 0–0.9 sound, 1.0–1.5 outer enamel, 1.5-2.0 inner enamel, 2.0+ dentine Device specifics: DIAGNOdent: conical tip (tip A) DIAGNOdent pen: cylindrical sapphire-fibre tip VistaProof: specific software (DBSWIN) that translates the rates of red and green fluorescence into numbers corresponding to lesion severity
Target condition and reference standard(s)	Category: histology Sequence of index test and reference standard: index test then reference standard Training of examiner: 2 trained examiners Blinding to index test: unclear Multiple tests: no Site selection: marked on photographs then sectioned teeth Target condition: sound, inner/outer enamel, inner/outer dentine
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0
Comparative	
Notes	

Diniz 2012 *(Continued)*

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Red fluorescence)			
DOMAIN 3: Reference Standard			

Diniz 2012 *(Continued)*

Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Diniz 2019

Study characteristics	
Patient Sampling	Method of sampling: unclear Included conditions: unclear on severity of lesions "with varying conditions from sound to that of different stages of carious lesion" Teeth: primary molars Sealants: excluded Surface: occlusal
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported Country: Brazil Setting: extracted teeth Number of participants/teeth/sites: 88 teeth Prevalence: enamel 0.74, dentine 0.63
Index tests	Category of test: DIAGNOdent, DIAGNOdent pen, QLF (QLF Inspektor Pro; Inspektor Research System, Amsterdam, Netherlands), and Midwest Sequence of test(s): index tests (visual, DIAGNOdent, DIAGNOdent pen, QLF, and MidWest) followed by reference standard

Diniz 2019 *(Continued)*

	Examiner training and calibration: 1 trained and experienced examiner Teeth cleaning prior to examination: rinsed with 3in1 syringe Tooth drying prior to examination: dried for DIAGNOdent, DIAGNOdent pen and QLF; but kept moist for Midwest Threshold applied: calculated within study for DIAGNOdent, DIAGNOdent pen, and QLF DIAGNOdent: 0–4 sound, 5-23 enamel, 24+ dentine DIAGNOdent pen: 0–3 sound, 4–19 enamel, 20+ dentine Midwest: green light/no beep - sound, red light slow/moderate beep - enamel, red light/fast beep - 3+ dentine QLF: 0-7.4 sound, 7.5-13.8 enamel, 13.9+ dentine Device specifics: using Inspektor Pro Software parameter ΔF (percentage of green fluorescence radiance loss) was recorded
Target condition and reference standard(s)	Category: histology Sequence of index test and reference standard: index test then reference standard Training of examiner: not reported Blinding to index test: unclear Multiple tests: no Site selection: marked on photographs then sectioned teeth Target condition: sound, inner/outer enamel, inner/outer dentine
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		

Diniz 2019 *(Continued)*

Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			

Diniz 2019 *(Continued)*

Was there an appropriate interval between index test and reference standard?	Yes	
Did all patients receive the same reference standard?	Yes	
Were all patients included in the analysis?	Yes	
Could the patient flow have introduced bias?		Low risk

Duruturk 2011

Study characteristics	
Patient Sampling	Method of sampling: selected "Teeth in which neither enamel nor dentin caries cavities were detected by visual or radiographic examination were measured using DIAGNOdent" Included conditions: no cavitation and early lesions Teeth: primary molars Sealants: no Surface: occlusal
Patient characteristics and setting	Age: 6 to 7 years Sex: not reported Ethnicity: not reported Country: Turkey Setting: attending pedodontic clinic Number of participants/teeth/sites: 307 participants/505 teeth/748 sites Prevalence: enamel 0.36, dentine not reported
Index tests	Category of test: DIAGNOdent Sequence of test(s): radiograph, visual, and DIAGNOdent followed by reference standard Examiner training and calibration: trained and calibrated Teeth cleaning prior to examination: professionally cleaned Tooth drying prior to examination: air spray 2 seconds Threshold applied: 0-14 sound, 15-20 enamel, 21+ dentine Device specifics: not reported
Target condition and reference standard(s)	Category: radiograph and visual combined Sequence of index test and reference standard: index test then reference standard Training of examiner: experienced

Duruturk 2011 *(Continued)*

	Blinding to index test: yes Multiple tests: combined test Site selection: teeth were drawn to aid examiners Target condition: sound, inner/outer enamel, inner/outer dentine		
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0		
Comparative			
Notes			
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	No		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			

Duruturk 2011 *(Continued)*

DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	No		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	No		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	No		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Unclear risk	

El-Housseiny 2001

Study characteristics	
Patient Sampling	Method of sampling: selected Included conditions: no cavitation and early lesions Teeth: permanent premolars and molars Sealants: no Surface: occlusal
Patient characteristics and setting	Age: not reported

El-Housseiny 2001 *(Continued)*

	Sex: not reported Ethnicity: not reported Country: Saudi Arabia Setting: extracted teeth Number of participants/teeth/sites: 46 teeth Prevalence: not reported
Index tests	Category of test: DIAGNOdent Sequence of test(s): index tests (visual and DIAGNOdent) followed by reference standard Examiner training and calibration: unclear Teeth cleaning prior to examination: pumice and rubber cups Tooth drying prior to examination: yes Threshold applied: 0-9 sound, 10-17 enamel, 18+ dentine Device specifics: tip A
Target condition and reference standard(s)	Category: histology Sequence of index test and reference standard: index test then reference standard Training of examiner: not reported Blinding to index test: unclear Multiple tests: no Site selection: sectioned teeth Target condition: sound, enamel, dentine
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns

El-Housseiny 2001 *(Continued)*

DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		

El-Housseiny 2001 *(Continued)*

Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Feng 2005

Study characteristics	
Patient Sampling	Method of sampling: unclear Included conditions: non-cavitated Teeth: unclear Sealants: no Surface: occlusal
Patient characteristics and setting	Age: 12 to 13 years Sex: 169 male, 131 female Ethnicity: not reported Country: China Setting: school based Number of participants/teeth/sites: 1732 teeth/300 participants Prevalence: 0.21 enamel
Index tests	Category of test: QLF Inspektor Research System BV, Amsterdam, Netherlands Sequence of test(s): visual then QLF and digital photo Examiner training and calibration: not reported Teeth cleaning prior to examination: professionally cleaned Tooth drying prior to examination: yes, with high pressure air (triple syringe) for 30 seconds

Feng 2005 *(Continued)*

	Threshold applied: "for QLF photos, upper anterior teeth that had decreased fluorescence in the cervical area were diagnosed as demineralization" Device specifics: none reported
Target condition and reference standard(s)	Category: visual, also digital photographs Sequence of index test and reference standard: unclear Training of examiner: not reported Blinding to index test: unclear Multiple tests: no Site selection: unclear Target condition: white spot lesions
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0
Comparative	
Notes	Translation completed by a Cochrane author, data extracted with visual as reference standard

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	No		
If a threshold was used, was it pre-specified?	Unclear		

Feng 2005 *(Continued)*

If multiple tests were applied were different examiners used for each (in vivo)?			
Could the conduct or interpretation of the index test have introduced bias?		Unclear risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	No		
If a threshold was used, was it pre-specified?	Unclear		
If multiple tests were applied were different examiners used for each (in vivo)?			
Could the conduct or interpretation of the index test have introduced bias?		Unclear risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	No		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Ferreira 1998

Study characteristics	
Patient Sampling	Method of sampling: selected Included conditions: no cavitation and early lesions Teeth: permanent premolars Sealants: no Surface: occlusal
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported Country: USA Setting: extracted teeth Number of participants/teeth/sites: 150 teeth Prevalence: not reported
Index tests	Category of test: bespoke LF device, combined with dye enhancement Sequence of test(s): index tests (visual and LF) followed by reference standard Examiner training and calibration: unclear Teeth cleaning prior to examination: yes Tooth drying prior to examination: 5 seconds Threshold applied: colour assessed by 2 examiners Device specifics: not relevant as unique device
Target condition and reference standard(s)	Category: histology Sequence of index test and reference standard: index test then reference standard Training of examiner: not reported Blinding to index test: unclear Multiple tests: no Site selection: marked on teeth Target condition: yes or no if decalcification present or not
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: minimal

Ferreira 1998 *(Continued)*

	Participants receiving both tests but excluded from results: 0		
Comparative			
Notes			
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			Unclear
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern

Ferreira 1998 *(Continued)*

DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			High
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Ferreira 2008

Study characteristics	
Patient Sampling	Method of sampling: selected Included conditions: no cavitation and early lesions Teeth: primary and permanent, incisors and molars Sealants: no Surface: smooth
Patient characteristics and setting	Age: 7 to 12 years Sex: not reported Ethnicity: not reported Country: Brazil Setting: school based Number of participants/teeth/sites: 36 participants Prevalence: not reported
Index tests	Category of test: DIAGNOdent

Ferreira 2008 *(Continued)*

	Sequence of test(s): index tests (DIAGNOdent) followed by reference standard (visual) Examiner training and calibration: yes Teeth cleaning prior to examination: professionally cleaned Tooth drying prior to examination: air dried 15 seconds Threshold applied: 0-4 sound or outer enamel, 5-10 inner enamel,10+ dentine Device specifics: tip B
Target condition and reference standard(s)	Category: visual Sequence of index test and reference standard: reference standard then index test Training of examiner: unclear Blinding to index test: no Multiple tests: no Site selection: unclear Target condition: healthy, activity with intact surfaces, inactivity
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	

Ferreira 2008 *(Continued)*

Are there concerns that the included patients and setting do not match the review question?			Unclear
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?			
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?			
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	No		
Were the reference standard results interpreted without knowledge of the results of the index tests?	No		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			High
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		

Ferreira 2008 *(Continued)*

Did all patients receive the same reference standard?	Yes	
Were all patients included in the analysis?	Yes	
Could the patient flow have introduced bias?		Low risk

Francescut 2003

Study characteristics	
Patient Sampling	Method of sampling: randomly selected Included conditions: no cavitation and early lesions; "They were macroscopically intact to the naked eye" Teeth: primary and permanent molars Sealants: no Surface: occlusal
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported Country: Switzerland Setting: extracted teeth Number of participants/teeth/sites: 190 teeth Prevalence: 0.18 dentine
Index tests	Category of test: DIAGNOdent Sequence of test(s): index tests (visual and DIAGNOdent) followed by reference standard Examiner training and calibration: yes Teeth cleaning prior to examination: calculus removed Tooth drying prior to examination: air dried 2 seconds Threshold applied: D2 > 14, D3 > 112 (D1 combined with sound) calculated within study, "For Diagnodent, the best cut-offs were set at a value in which the maximal sensitivity and specificity were obtained" Device specifics: not reported
Target condition and reference standard(s)	Category: histology Sequence of index test and reference standard: index test then reference standard Training of examiner: not reported Blinding to index test: unclear

Francescut 2003 *(Continued)*

	Multiple tests: no Site selection: sectioned teeth Target condition: caries: enamel outer, enamel inner, outer dentine, inner dentine		
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0		
Comparative			
Notes	Data only available at the dentine level due to the combining of sound and non-cavitated lesions		

Methodological quality

Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		
If multiple tests were applied were different examiners used for each (in vivo)?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			High
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			

Francescut 2003 *(Continued)*

Were the index test results interpreted without knowledge of the results of the reference standard?			
If a threshold was used, was it pre-specified?			
If multiple tests were applied were different examiners used for each (in vivo)?			
Could the conduct or interpretation of the index test have introduced bias?			
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Fung 2004

Study characteristics	
Patient Sampling	Method of sampling: selected/unclear Included conditions: no cavitation and early lesions Teeth: permanent molars Sealants: unclear Surface: occlusal
Patient characteristics and setting	Age: not reported Sex: not reported

Fung 2004 *(Continued)*

	Ethnicity: not reported Country: Australia Setting: extracted teeth Number of participants/teeth/sites: 25 teeth Prevalence: unclear
Index tests	Category of test: DIAGNOdent Sequence of test(s): index tests (visual and DIAGNOdent) followed by reference standard Examiner training and calibration: yes Teeth cleaning prior to examination: pumice and slurry Tooth drying prior to examination: air dried Threshold applied: "a conservative cut-off limit of 30 was used" assumed for dentine threshold Device specifics: tip A
Target condition and reference standard(s)	Category: histology Sequence of index test and reference standard: index test then reference standard Training of examiner: not reported Blinding to index test: unclear Multiple tests: no Site selection: sectioned teeth Target condition: caries: "Caries in either enamel or dentine was diagnosed"
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 20 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			

Fung 2004 *(Continued)*

Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?			
If a threshold was used, was it pre-specified?			
If multiple tests were applied were different examiners used for each (in vivo)?			
Could the conduct or interpretation of the index test have introduced bias?			
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	

Fung 2004 *(Continued)*

Are there concerns that the target condition as defined by the reference standard does not match the question?		Low concern
DOMAIN 4: Flow and Timing		
Was there an appropriate interval between index test and reference standard?	Yes	
Did all patients receive the same reference standard?	Yes	
Were all patients included in the analysis?	No	
Could the patient flow have introduced bias?		High risk

Ghaname 2010

Study characteristics	
Patient Sampling	Method of sampling: selected Included conditions: no cavitation and early lesions (includes up to ICDAS 3) Teeth: permanent molars and premolars Sealants: unclear Surface: occlusal
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported Country: US Setting: extracted teeth Number of participants/teeth/sites: 103 sites/teeth Prevalence: dentine 0.29
Index tests	Category of test: DIAGNOdent Sequence of test(s): index tests (DIAGNOdent) followed by reference standard Examiner training and calibration: yes Teeth cleaning prior to examination: soft tissues removed with hand instruments Tooth drying prior to examination: "briefly dried" Threshold applied: > 20 assumed for dentine threshold Device specifics: not reported

Ghaname 2010 *(Continued)*

Target condition and reference standard(s)	Category: excavation with "Lesion Volume and Extension Determination" Sequence of index test and reference standard: index test then reference standard Training of examiner: experienced Blinding to index test: yes Multiple tests: combined test Site selection: opened all occlusal fissures Target condition: dentinal or no dentinal caries
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?			
Could the conduct or interpretation of the index test have introduced bias?		Low risk	

Ghaname 2010 *(Continued)*

Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?			
If a threshold was used, was it pre-specified?			
If multiple tests were applied were different examiners used for each (in vivo)?			
Could the conduct or interpretation of the index test have introduced bias?			
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			High
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Goel 2009

Study characteristics	
Patient Sampling	Method of sampling: selected Included conditions: no cavitation and early enamel lesions

Goel 2009 *(Continued)*

	Teeth: first and second molars Sealants: no Surface: occlusal
Patient characteristics and setting	Age: 8 to 12 years Sex: not reported Ethnicity: not reported Country: India Setting: index test performed in a clinical setting prior to extraction Number of participants/teeth/sites: 84 teeth/83 sites Prevalence: enamel 0.54, dentine 0.43
Index tests	Category of test: DIAGNOdent Sequence of test(s): index tests (radiograph, visual, and DIAGNOdent) performed prior to reference standard Examiner training and calibration: not reported Teeth cleaning prior to examination: calculus removed with scaler Tooth drying prior to examination: air dried 5 seconds Threshold applied: 0-5 sound, 6-14 outer enamel, 15-20 inner enamel, 21+ dentinal Device specifics: none reported
Target condition and reference standard(s)	Category: histology Sequence of index test and reference standard: index test then reference standard Training of examiner: not reported Blinding to index test: unclear Multiple tests: no Site selection: sectioned teeth Target condition: sound, inner/outer enamel, inner/outer dentine
Flow and timing	Participants with index test but no reference standard: 1 Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0
Comparative	

Goel 2009 *(Continued)*

Notes

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?			
If a threshold was used, was it pre-specified?			
If multiple tests were applied were different examiners used for each (in vivo)?			
Could the conduct or interpretation of the index test have introduced bias?			
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			
DOMAIN 3: Reference Standard			

Goel 2009 *(Continued)*

Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	No		
Could the patient flow have introduced bias?		Low risk	

Graye 2012

Study characteristics	
Patient Sampling	Method of sampling: selected Included conditions: unclear Teeth: third molars Sealants: no Surface: occlusal
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported Country: USA Setting: extracted teeth Number of participants/teeth/sites: 41 teeth Prevalence: enamel 0.90, dentine 0.46
Index tests	Category of test: Spectra Sequence of test(s): index tests (radiograph, visual, and Spectra) performed prior to reference standard Examiner training and calibration: not reported Teeth cleaning prior to examination: debris removed

Graye 2012 *(Continued)*

	Tooth drying prior to examination: yes Threshold applied: green 0-1 sound, blue 1-1.5 outer enamel, red 1.5-2 inner enamel, orange/yellow 2+ dentine Device specifics: uses accompanying software
Target condition and reference standard(s)	Category: histology Sequence of index test and reference standard: index test then reference standard Training of examiner: not reported Blinding to index test: unclear Multiple tests: no Site selection: sectioned teeth Target condition: sound, inner/outer enamel, inner/outer dentine
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		

Graye 2012 *(Continued)*

If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Heinrich-Weltzien 2003

Study characteristics	
Patient Sampling	Method of sampling: selected Included conditions: unclear Teeth: first and second molars, permanent Sealants: no Surface: occlusal
Patient characteristics and setting	Age: mean age 19.2 years Sex: not reported Ethnicity: not reported Country: Germany Setting: general dental setting Number of participants/teeth/sites: 94 participants/248 teeth/sites Prevalence: enamel 0.90, dentine 0.85
Index tests	Category of test: DIAGNOdent Sequence of test(s): index tests (visual and DIAGNOdent) performed prior to reference standard Examiner training and calibration: experienced Teeth cleaning prior to examination: professionally cleaned Tooth drying prior to examination: airflow device Threshold applied: defined and investigated within study Device specifics: conical probe A
Target condition and reference standard(s)	Category: excavation Sequence of index test and reference standard: index test then reference standard Training of examiner: not reported Blinding to index test: unclear Multiple tests: no Site selection: no biopsy on sound lesions, so assumed visual examination used as reference standard for those surfaces Target condition: sound, enamel, dentinal lesions
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: minimal

Heinrich-Weltzien 2003 *(Continued)*

	Participants receiving both tests but excluded from results: 0		
Comparative			
Notes			
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	

Heinrich-Weltzien 2003 *(Continued)*

Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	No		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	No		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		High risk	

Hibst 2001

Study characteristics	
Patient Sampling	Method of sampling: unclear Included conditions: not reported Teeth: not reported Sealants: no Surface: occlusal
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported Country: Germany Setting: extracted teeth Number of participants/teeth/sites: 240 participants/332 teeth/sites Prevalence: not reported
Index tests	Category of test: DIAGNOdent

Hibst 2001 *(Continued)*

	Sequence of test(s): index tests performed prior to reference standard Examiner training and calibration: unclear, completed by a dental professional Teeth cleaning prior to examination: not reported Tooth drying prior to examination: not reported Threshold applied: sound < 14, enamel 14-20, dentine > 20 Device specifics: not reported
Target condition and reference standard(s)	Category: excavation Sequence of index test and reference standard: index test then reference standard Training of examiner: not reported Blinding to index test: unclear Multiple tests: no Site selection: when tooth required opening Target condition: sound, enamel, dentinal lesions
Flow and timing	Participants with index test but no reference standard: not reported but some will not have received the excavation reference standard Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			High

Hibst 2001 *(Continued)*

DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	No		
Were the reference standard results interpreted without knowledge of the results of the index tests?	No		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	No		

Hibst 2001 *(Continued)*

Were all patients included in the analysis?	Yes
Could the patient flow have introduced bias?	High risk

Huth 2008

Study characteristics	
Patient Sampling	Method of sampling: randomised Included conditions: no cavitation and early lesions Teeth: permanent molars Sealants: no Surface: occlusal
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported Country: Switzerland Setting: dental hospital Number of participants/teeth/sites: 120 participants Prevalence: not reported
Index tests	Category of test: DIAGNOdent pen Sequence of test(s): index tests (visual, radiograph, DIAGNOdent pen) performed prior to reference standard Examiner training and calibration: unclear, completed by a dental professional Teeth cleaning prior to examination: yes Tooth drying prior to examination: yes Threshold applied: calculated in study, multiple thresholds investigated within study Device specifics: not reported
Target condition and reference standard(s)	Category: excavation or visual/radiograph with follow-up Sequence of index test and reference standard: index test then reference standard Training of examiner: experienced Blinding to index test: unclear Multiple tests: yes

Huth 2008 *(Continued)*

	Site selection: unclear which site was investigated with which test Target condition: sound, enamel, dentinal lesions		
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0		
Comparative			
Notes	Data not useable as reported the mean for DIAGNOdent readings		

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		
If multiple tests were applied were different examiners used for each (in vivo)?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			

Huth 2008 *(Continued)*

Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		
If multiple tests were applied were different examiners used for each (in vivo)?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	No		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	No		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		High risk	

Huth 2010

Study characteristics	
Patient Sampling	Method of sampling: randomised Included conditions: no cavitation and early lesions Teeth: permanent molars Sealants: no Surface: occlusal
Patient characteristics and setting	Age: not reported Sex: not reported

Huth 2010 *(Continued)*

	Ethnicity: not reported Country: Switzerland Setting: dental hospital Number of participants/teeth/sites: 117 participants Prevalence: enamel 0.66, dentine 0.37
Index tests	Category of test: DIAGNOdent pen Sequence of test(s): index tests (visual, radiograph, DIAGNOdent pen) performed prior to reference standard Examiner training and calibration: unclear, completed by a dental professional Teeth cleaning prior to examination: yes Tooth drying prior to examination: yes Threshold applied: calculated in study, multiple thresholds investigated within study (cut-off at D1 level = 7) Device specifics: not reported
Target condition and reference standard(s)	Category: excavation or visual/radiograph with follow-up Sequence of index test and reference standard: index test then reference standard Training of examiner: experienced, same examiners as index tests or aware of the results of the index test Blinding to index test: unclear Multiple tests: yes Site selection: unclear which site was investigated with which test Target condition: sound, enamel, dentinal lesions
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0
Comparative	
Notes	Data used for the in vivo level, from table 3 (D0 versus D1-4)

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			

Huth 2010 *(Continued)*

Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		
If multiple tests were applied were different examiners used for each (in vivo)?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		
If multiple tests were applied were different examiners used for each (in vivo)?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	No		
Were the reference standard results interpreted without knowledge of the results of the index tests?	No		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	

Huth 2010 *(Continued)*

Are there concerns that the target condition as defined by the reference standard does not match the question?		Low concern
DOMAIN 4: Flow and Timing		
Was there an appropriate interval between index test and reference standard?	Yes	
Did all patients receive the same reference standard?	No	
Were all patients included in the analysis?	Yes	
Could the patient flow have introduced bias?		High risk

Iranzo-Cortes 2017

Study characteristics	
Patient Sampling	Method of sampling: selected - "teeth extracted for orthodontic or periodontal reasons was selected" Included conditions: "healthy or present incipient caries lesions but those with large cavitated lesions or filled surfaces were excluded" Teeth: permanent premolars and molars Sealants: unclear Surface: occlusal
Patient characteristics and setting	Age: 18 to 55 years Sex: not reported Ethnicity: not reported Country: Spain Setting: extracted teeth Number of participants/teeth/sites: 65 teeth Prevalence: 0.77 enamel, 0.17 dentine
Index tests	Category of test: DIAGNOdent Sequence of test(s): index tests performed (visual then DIAGNOdent) prior to reference standard Examiner training and calibration: 35 teeth used for calibration Teeth cleaning prior to examination: calculus and residues were removed from the selected teeth, using a KAVO Sonic Flex Tooth drying prior to examination: triple air syringe was used to dry teeth

Iranzo-Cortes 2017 *(Continued)*

	Threshold applied: sound < 14, enamel 14-29, dentine > 30 Device specifics: not reported
Target condition and reference standard(s)	Category: histology Sequence of index test and reference standard: index test then reference standard Training of examiner: not reported Blinding to index test: unclear Multiple tests: no Site selection: ss marked prior to index test, then sectioned teeth Target condition: sound, inner/outer enamel, inner/outer dentine
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	Unclear		

Iranzo-Cortes 2017 *(Continued)*

Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Jablonski-Momeni 2011

Study characteristics

Jablonski-Momeni 2011 *(Continued)*

Patient Sampling	Method of sampling: selected Included conditions: unclear Teeth: permanent molars and premolars Sealants: not reported Surface: occlusal
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported Country: Germany Setting: extracted teeth Number of participants/teeth/sites: 53 teeth/99 sites Prevalence: enamel 0.76, dentine 0.23
Index tests	Category of test: VistaProof Sequence of test(s): index tests (visual and VistaProof) performed prior to reference standard Examiner training and calibration: experienced trained dentist Teeth cleaning prior to examination: yes, method not reported Tooth drying prior to examination: not reported Threshold applied: sound 0-0.9, initial enamel 0.9-1.5, deep enamel 1.5-2.0, dentine 2+ Device specifics: a long distance space was used, DBSWIN software used for analysis
Target condition and reference standard(s)	Category: histology Sequence of index test and reference standard: index test then reference standard Training of examiner: not reported Blinding to index test: unclear Multiple tests: no Site selection: sectioned teeth Target condition: sound, outer enamel, inner enamel, dentine
Flow and timing	Participants with index test but no reference standard: 1 Participants with reference standard but no index test: 0 Time interval between tests: minimal

Jablonski-Momeni 2011 *(Continued)*

	Participants receiving both tests but excluded from results: 0		
Comparative			
Notes			
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	

Jablonski-Momeni 2011 *(Continued)*

Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Red fluorescence)			
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Jablonski-Momeni 2012

Study characteristics	
Patient Sampling	Method of sampling: selected Included conditions: no cavitation and early lesions (low number of potentially dentinal lesions) Teeth: permanent molars Sealants: not reported Surface: occlusal
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported Country: Germany Setting: extracted teeth Number of participants/teeth/sites: 36 teeth/82 sites Prevalence: enamel 0.72, dentine 0.21

Jablonski-Momeni 2012 *(Continued)*

Index tests	Category of test: DIAGNOdent pen and VistaProof Sequence of test(s): index tests (visual then DIAGNOdent and VistaProof) performed prior to reference standard Examiner training and calibration: 2 trained examiners Teeth cleaning prior to examination: yes, method not reported Tooth drying prior to examination: not reported Threshold applied: DIAGNOdent pen: 0-6 sound; 6-13 enamel caries; 13-17 enamel caries to EDJ; > 17 dentine caries VistaProof: 0.0-0.9 sound; 0.9-1.5 enamel caries; 1.5-2.0 enamel caries to EDJ; > 2.0 dentine caries Device specifics: DIAGNOdent pen: tip A VistaProof: long-distance spacer, DBSWIN software used to analyse results
Target condition and reference standard(s)	Category: histology Sequence of index test and reference standard: index test then reference standard Training of examiner: not reported Blinding to index test: unclear Multiple tests: no Site selection: sectioned teeth Target condition: sound, outer/inner enamel, outer/inner dentine
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		

Jablonski-Momeni 2012 *(Continued)*

Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Red fluorescence)			
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	

Jablonski-Momeni 2012 *(Continued)*

Are there concerns that the target condition as defined by the reference standard does not match the question?		Low concern
DOMAIN 4: Flow and Timing		
Was there an appropriate interval between index test and reference standard?	Yes	
Did all patients receive the same reference standard?	Yes	
Were all patients included in the analysis?	Yes	
Could the patient flow have introduced bias?		Low risk

Jablonski-Momeni 2012a

Study characteristics	
Patient Sampling	Method of sampling: unclear Included conditions: no cavitation and early lesions Teeth: permanent premolars and molars Sealants: not reported Surface: occlusal; "permanent posterior teeth without occlusal restorations"
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported Country: Germany Setting: extracted teeth Number of participants/teeth/sites: 36 teeth/82 sites Prevalence: enamel 0.84, dentine 0.48
Index tests	Category of test: DIAGNOdent and VistaCam iX (using fluorescence) Sequence of test(s): index tests (visual then DIAGNOdent and VistaCam) performed prior to reference standard Examiner training and calibration: 2 examiners, "doctoral student calibrated by an experienced investigator" Teeth cleaning prior to examination: yes, method not reported Tooth drying prior to examination: not reported Threshold applied: DIAGNOdent: 0-7 sound; 8-24 enamel caries; > 25 dentine caries

Jablonski-Momeni 2012a *(Continued)*

	VistaCam: 0.0-0.9 sound; 0.9-2.0 enamel; > 2.0 dentine caries (manufacturers thresholds) Device specifics: DIAGNOdent pen: tip A VistaProof: long-distance spacer
Target condition and reference standard(s)	Category: excavation Sequence of index test and reference standard: index test then reference standard Training of examiner: not reported Blinding to index test: unclear Multiple tests: no Site selection: all teeth opened with rotating instrument Target condition: sound, enamel, dentine
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 4, "While using the FC device, 4 investigation sites could not be assessed due to technical problems" Time interval between tests: minimal Participants receiving both tests but excluded from results: 0
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		

Jablonski-Momeni 2012a *(Continued)*

If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		Unclear risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		Unclear risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Red fluorescence)			
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	No		
Could the patient flow have introduced bias?		Low risk	

Jablonski-Momeni 2014

Study characteristics	
Patient Sampling	Method of sampling: selected Included conditions: not clearly stated in the recruitment section, results report acceptable level of dentinal lesions Teeth: permanent molars and premolars Sealants: not reported Surface: occlusal
Patient characteristics and setting	Age: minimum age of 18 years, mean 27.4 Sex: 10 male, 16 female Ethnicity: not reported Country: Germany Setting: assumed to be a clinical setting as the aim was to determine which surfaces should be restored Number of participants/teeth/sites: 26 teeth/306 sites Prevalence: enamel 0.17, dentine 0.12
Index tests	Category of test: VistaProof Sequence of test(s): unclear on the sequence of tests, reported as visual first then VistaProof followed by radiograph and excavation where propriae Examiner training and calibration: 2 trained examiners Teeth cleaning prior to examination: yes, cleaned and air-dried using a triplex syringe Tooth drying prior to examination: as above Threshold applied: 0–0.9 sound; 1.0–1.4 early stage of enamel lesion; 1.5–1.9 deep enamel lesion; 2.0–2.4 dentine caries; and > 2.4 deep dentine caries Device specifics: specific software used for analysis, "Sound enamel and carious lesions are visualised in colour and numerically (on a scale from 0 to 4)"
Target condition and reference standard(s)	Category: visual (ICDAS) for all surfaces, where appropriate radiographs and excavation where applied Sequence of index test and reference standard: it seems the index test was performed after visual examination and before radiographs, so index may have influenced decision Training of examiner: 1 experienced examiner Blinding to index test: no Multiple tests: yes Site selection: all selected occlusal surfaces

Jablonski-Momeni 2014 *(Continued)*

	Target condition: ICDAS categories: 0 = sound; 1 = first visible sign of non-cavitated lesion seen only when the tooth is dry; 2 = visible non-cavitated lesion seen when wet and dry; 3 = microcavitation in enamel; code 4 = non-cavitated lesion extending into dentine seen as an undermining shadow; code 5 = small cavitated lesion with visible dentine: less than 50% of surface; and code 6 = large cavitated lesion with visible dentine
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?			
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			

Jablonski-Momeni 2014 *(Continued)*

Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?			
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Red fluorescence)			
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	No		
Were the reference standard results interpreted without knowledge of the results of the index tests?	No		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	No		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		High risk	

Jablonski-Momeni 2016

Study characteristics	
Patient Sampling	Method of sampling: unclear Included conditions: non-cavitated and early lesions (ICDAS 0-2) Teeth: primary (this entry is for primary) and permanent Sealants: not reported Surface: occlusal

Jablonski-Momeni 2016 *(Continued)*

Patient characteristics and setting	Age: 5 to 12 years, mean age 9.1 Sex: 17 male, 18 female Ethnicity: not reported Country: Germany Setting: "recruited in a dental office" Number of participants/teeth/sites: 35 participants/205 primary, 214 permanent teeth Prevalence: primary: enamel 0.18, dentine 0 permanent: enamel 0.35, dentine 0
Index tests	Category of test: VistaProof Sequence of test(s): visual prior to VistaProof, so reference standard then index test Examiner training and calibration: not reported Teeth cleaning prior to examination: yes, rotating brush and paste, then rinsed with a 3 in 1 syringe Tooth drying prior to examination: unclear Threshold applied: 0–1.2 sound; 1.3–1.5 enamel caries; > 1.5 dentine caries Device specifics: "Each image was analyzed by the specific software (DBSWIN, Durr Dental)"
Target condition and reference standard(s)	Category: visual (ICDAS) for all surfaces Sequence of index test and reference standard: the index test was performed after visual examination Training of examiner: unclear Blinding to index test: not reported Multiple tests: no Site selection: all selected occlusal surfaces Target condition: ICDAS categories: 0 = sound; 1 = first visible sign of non-cavitated lesion seen only when the tooth is dry; 2 = visible non-cavitated lesion seen when wet and dry; 3 = microcavitation in enamel; code 4 = non-cavitated lesion extending into dentine seen as an undermining shadow; code 5 = small cavitated lesion with visible dentine: less than 50% of surface; and code 6 = large cavitated lesion with visible dentine
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: minimal

Jablonski-Momeni 2016 *(Continued)*

	Participants receiving both tests but excluded from results: 13, "Thirteen teeth were unable to be monitored for 1 year (due to restorative treatment or extraction)"		
Comparative			
Notes			
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?			
Could the conduct or interpretation of the index test have introduced bias?		Unclear risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?			
Could the conduct or interpretation of the index test have introduced bias?		Unclear risk	

Jablonski-Momeni 2016 *(Continued)*

Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Red fluorescence)			
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	No		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			High
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	No		
Could the patient flow have introduced bias?		Low risk	

Jeon 2004

Study characteristics	
Patient Sampling	Method of sampling: selected Included conditions: no cavitation and early lesions Teeth: permanent/primary premolars and molars Sealants: unclear Surface: occlusal and smooth
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported Country: Canada Setting: extracted teeth Number of participants/teeth/sites: 52 teeth/332 sites (104 healthy points, 176 occlusal fissures, 52 healthy points on the smooth surface)

Jeon 2004 *(Continued)*

	Prevalence: enamel level not reported, dentine 0.16 (from the DIAGNOdent results in table 3)
Index tests	Category of test: DIAGNOdent (completed on 131 sites - Table 3 in paper) Sequence of test(s): index tests performed (visual, radiograph then DIAGNOdent) prior to reference standard Examiner training and calibration: not reported Teeth cleaning prior to examination: not reported Tooth drying prior to examination: not reported Threshold applied: 0-4 sound or outer enamel, 4.01-10 inner enamel, 10.01-18 outer dentine, 18.01+ inner dentine Device specifics: not reported
Target condition and reference standard(s)	Category: histology Sequence of index test and reference standard: index test then reference standard Training of examiner: not reported Blinding to index test: unclear Multiple tests: no Site selection: marked on a photograph prior to index test, then sectioned teeth Target condition: sound, inner/outer enamel, inner/outer dentine
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: differs, some examiners did not assess all sites Time interval between tests: minimal Participants receiving both tests but excluded from results: 0
Comparative	
Notes	Study also assesses frequency-domain photothermal radiometry and frequency-domain luminescence

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		

Jeon 2004 *(Continued)*

Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			Unclear
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	No		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	No		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			

Jeon 2004 *(Continued)*

Was there an appropriate interval between index test and reference standard?	Yes	
Did all patients receive the same reference standard?	No	
Were all patients included in the analysis?	No	
Could the patient flow have introduced bias?		High risk

Jung 2018

Study characteristics	
Patient Sampling	Method of sampling: consecutive Included conditions: no cavitation and early lesions Teeth: permanent Sealants: no Surface: occlusal
Patient characteristics and setting	Age: minimum age of 18 years Sex: not reported Ethnicity: not reported Country: South Korea Setting: extracted teeth Number of participants/teeth/sites: 94 participants/791 teeth Prevalence: enamel 0.47, dentine 0.14
Index tests	Category of test: QLF images Sequence of test(s): index tests performed (visual then QLF) prior to histology Examiner training and calibration: yes - single calibrated examiner Teeth cleaning prior to examination: professionally by therapists Tooth drying prior to examination: "sufficient drying" Threshold applied: sound, initial caries, enamel caries, dentine Device specifics: not reported clearly - QS-Occlusal software algorithm was used to determine the levels of disease D1 = 0/1, D2 = 1/2, D3 = 3/4
Target condition and reference standard(s)	Category: visual ICDAS classification Sequence of index test and reference standard: visual examination completed prior to QLF with histology following Training of examiner: not reported

Jung 2018 *(Continued)*

	Blinding to index test: unclear - examiner not blinded between visual and QLF, although 2 weeks passed between assessments Multiple tests: yes Site selection: marked on a photograph prior to index test, then sectioned teeth Target condition: ICDAS codes
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: lack of clarity on true reference standard Time interval between tests: 2 weeks Participants receiving both tests but excluded from results: 0
Comparative	
Notes	D1 threshold used as labelled in table 4 as 0 versus 1-4

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	No		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			High
DOMAIN 2: Index Test (Green fluorescence)			

Jung 2018 *(Continued)*

Were the index test results interpreted without knowledge of the results of the reference standard?	No		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			High
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	No		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Unclear		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Unclear risk	

Kavvadia 2008

Study characteristics	
Patient Sampling	Method of sampling: selected Included conditions: no cavitation and early lesions Teeth: primary molars Sealants: not reported Surface: occlusal

Kavvadia 2008 *(Continued)*

Patient characteristics and setting	Age: 3 to 13 years, mean 5.94 Sex: 26 male, 21 female Ethnicity: not reported Country: Greece Setting: dental hospital Number of participants/teeth/sites: 47 participants/130 teeth/405 sites Prevalence: enamel 0.38, dentine 0.18
Index tests	Category of test: DIAGNOdent Sequence of test(s): index tests performed (visual, radiograph, then DIAGNOdent) prior to reference standard Examiner training and calibration: calibrated Teeth cleaning prior to examination: rubber cup and pumice Tooth drying prior to examination: air dried Threshold applied: calculated within study using Speaman's correlation coefficient: 0-9 sound, 10-42 enamel, 30-99 dentine Device specifics: tip A
Target condition and reference standard(s)	Category: excavation following results of visual/radiograph examination Sequence of index test and reference standard: index test then reference standard Training of examiner: not reported Blinding to index test: unclear Multiple tests: yes Site selection: evaluated pits or fissures Target condition: sound, enamel, dentinal lesions
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: lack of clarity application of reference standard Time interval between tests: minimal Participants receiving both tests but excluded from results: 0
Comparative	
Notes	Cannot include test data as reference standard only reported on carious teeth
Methodological quality	

Kavvadia 2008 *(Continued)*

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	No		

Kavvadia 2008 *(Continued)*

Were the reference standard results interpreted without knowledge of the results of the index tests?	No		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	No		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		High risk	

Kavvadia 2012

Study characteristics	
Patient Sampling	Method of sampling: selected Included conditions: no cavitation and early lesions Teeth: primary molars Sealants: no Surface: occlusal
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported Country: Greece Setting: extracted teeth Number of participants/teeth/sites: 47 participants/24 teeth/111 sites Prevalence: enamel 0.98, dentine 0.22
Index tests	Category of test: DIAGNOdent Sequence of test(s): index tests performed (visual, radiograph, then DIAGNOdent) prior to reference standard Examiner training and calibration: not reported Teeth cleaning prior to examination: brush and paste Tooth drying prior to examination: not reported

Kavvadia 2012 *(Continued)*

	Threshold applied: generated within study using Speaman's correlation coefficient: 0-2 sound, 3-39 enamel, 40-99 dentine Device specifics: tip A
Target condition and reference standard(s)	Category: histology Sequence of index test and reference standard: index test then reference standard Training of examiner: not reported Blinding to index test: unclear Multiple tests: no Site selection: sectioned teeth Target condition: sound, inner/outer enamel, inner/outer dentine
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests:minimal Participants receiving both tests but excluded from results: 0
Comparative	
Notes	Data not included as not possible to extract into a 2x2 table from table 5, 3 thresholds are reported and therefore unclear which results are appropriate for our 2 thresholds used

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		

Kavvadia 2012 *(Continued)*

If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Kesler 2003

Study characteristics	
Patient Sampling	Method of sampling: selected Included conditions: no cavitation and early lesions Teeth: primary molars Sealants: no Surface: occlusal
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported Country: Israel Setting: extracted teeth Number of participants/teeth/sites: 901 teeth Prevalence: enamel 0.67, dentine 0.19
Index tests	Category of test: fluorescence - Oliver 101 Sequence of test(s): index tests performed (visual, radiograph, then fluorescence) prior to reference standard Examiner training and calibration: not reported Teeth cleaning prior to examination: polished and cleaned Tooth drying prior to examination: not reported Threshold applied: calculated within study Device specifics: unclear
Target condition and reference standard(s)	Category: histology Sequence of index test and reference standard: index test then reference standard Training of examiner: not reported Blinding to index test: unclear Multiple tests: no Site selection: sectioned teeth Target condition: enamel or dentine caries
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: minimal

Kesler 2003 *(Continued)*

	Participants receiving both tests but excluded from results: 0		
Comparative			
Notes			
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		
If multiple tests were applied were different examiners used for each (in vivo)?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		
If multiple tests were applied were different examiners used for each (in vivo)?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern

Kesler 2003 *(Continued)*

DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Kim 2017

Study characteristics	
Patient Sampling	Method of sampling: not clearly reported Included conditions: severity of condition unclear, "subjects with 1 or more proximal caries surfaces detected visually or radiographically were included in the study", restorations were included Teeth: permanent molars and premolars Sealants: not reported Surface: approximal
Patient characteristics and setting	Age: 19 to 60 years Sex: 55% male Ethnicity: not reported Country: South Korea Setting: extracted teeth Number of participants/teeth/sites: 65 teeth/280 sites Prevalence: enamel 0.61, dentine 0.20

Kim 2017 *(Continued)*

Index tests	Category of test: QLF-Digital Biluminator 2+ (QLF-D, Inspektor Research Systems BV, Amsterdam, The Netherlands), 2 methods one using QA2 software, the second using fluorescence images interpreted by an examiner: "Normal white-light images and sequential fluorescence images were captured with a “live view” enabled full-frame sensor digital SLR camera" Sequence of test(s): visual then radiograph followed by QLF, radiograph was the reference standard Examiner training and calibration: 1 trained examiner completed all index tests and reference standard Teeth cleaning prior to examination: full-mouth scaling and polishing Tooth drying prior to examination: not reported Threshold applied: method used for fluorescence image method: shadow and no red fluorescence (Q0), an irregular dark shadow but no red fluorescence (Q1), faint red fluorescence limited to 1/3 of the buccolingual width (Q2), and strong red fluorescence over 1/3 of the buccolingual width (Q3)
Target condition and reference standard(s)	Category: radiograph Sequence of index test and reference standard: reference standard prior to index test Training of examiner: not reported, but experienced Blinding to index test: no Multiple tests: no Site selection: approximal surfaces Target condition: sound, outer/inner enamel, outer/inner dentine
Flow and timing	Participants with index test but no reference standard: 15 Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0
Comparative	
Notes	Data used for the fluorescence images method as the 2x2 figures were not available for the software method

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		

Kim 2017 *(Continued)*

Did the study avoid inappropriate exclusions?	Unclear		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	No		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	No		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	No		
Were the reference standard results interpreted without knowledge of the results of the index tests?	No		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	

Kim 2017 *(Continued)*

Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	No		
Could the patient flow have introduced bias?		High risk	

Ko 2015

Study characteristics	
Patient Sampling	Method of sampling: selected Included conditions: severity of condition unclear, "proximal surfaces with extensive cavities involving more than half of the proximal surface were excluded" Teeth: permanent molars and premolars Sealants: not reported Surface: approximal
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported Country: South Korea Setting: extracted teeth Number of participants/teeth/sites: 100 teeth (5 were damaged so only 95 reported in results) Prevalence: enamel 0.80, dentine 0.15
Index tests	Category of test: QLF-Digital Biluminator (QLF-D, Inspektor Research Systems BV, Amsterdam, The Netherlands), using proprietary software (C3 v 1.16); "Pairs were formed with marginal ridges in contact to simulate the oral relationship" Sequence of test(s): visual then radiograph followed by QLF Examiner training and calibration: 1 calibrated dentist Teeth cleaning prior to examination: cleaned of all soft tissues Tooth drying prior to examination: dried with cotton wool

Ko 2015 *(Continued)*

	Threshold applied: calculated within study, sound < -13.8, enamel -13.8 to -28.3, dentine > -28.3 Device specifics: shutter speed 1-20 seconds, aperture 13.0, ISO speed 1600, 10 cm between specimen and the device
Target condition and reference standard(s)	Category: histology Sequence of index test and reference standard: index test then reference standard Training of examiner: not reported Blinding to index test: unclear Multiple tests: no Site selection: sectioned teeth Target condition: "enamel demineralization or a narrow surface zone of opacity", enamel or outer/inner dentine caries
Flow and timing	Participants with index test but no reference standard: 5 Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		

Ko 2015 *(Continued)*

If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	No		
Could the patient flow have introduced bias?		Low risk	

Kockanat 2017

Study characteristics	
Patient Sampling	Method of sampling: unclear Included conditions: non-cavitated; "occlusal surfaces of the teeth had minimal macroscopic destruction" Teeth: primary molars Sealants: unclear Surface: occlusal
Patient characteristics and setting	Age: 9 to 12 years Sex: not reported Ethnicity: not reported Country: Turkey Setting: in vivo study conducted in dental hospital, followed by in vitro after extraction Number of participants/teeth/sites: 120 teeth (144 teeth were examined and measurements made with caries detection devices, but 120 of the 144 teeth were reported; due to inconsistencies in caries measurement results), clarification provided by study author Prevalence: enamel 0.78, dentine 0.32
Index tests	Category of test: DIAGNOdent pen and Sopro camera Sequence of test(s): visual, SoproLife, DIAGNOdent pen then CarieScan PRO Examiner training and calibration: unclear, 2 independent examiners Teeth cleaning prior to examination: plaque removed, washed without pumice Tooth drying prior to examination: air water spray, dried again for 5 seconds prior to DD Threshold applied: DIAGNOdent pen: 0-13 sound, 14-29 enamel, 30+ dentine Sopro camera: (0) no visible radiolucency; (1) radiolucency in the enamel; (2) radiolucency in the dentine, involving the surface or the outer third of the dentine, and (3) radiolucency in the dentine, involving the inner third of the dentine Device specifics: cylinder sapphire tip for DIAGNOdent pen, "The images were recorded to Sopro Imaging program and evaluated according to the criteria of Rechmann"
Target condition and reference standard(s)	Category: histology Sequence of index test and reference standard: index tests then reference standard Training of examiner: not reported

Kockanat 2017 *(Continued)*

	Blinding to index test: not reported Multiple tests: no Site selection: sectioned teeth Target condition: sound, outer half of enamel, inner half of enamel, outer half of dentine, deep dentine
Flow and timing	Participants with index test but no reference standard: 24 Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0
Comparative	
Notes	Data used for examiner 1, with the comparison of in vivo index test versus histology Study authors contacted for clarification of study data

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern

Kockanat 2017 *(Continued)*

DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Red fluorescence)			
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	No		
Could the patient flow have introduced bias?		High risk	

Kouchaji 2012

Study characteristics	
Patient Sampling	Method of sampling: unclear Included conditions: no cavitation or early lesions, "The study used first permanent molars with and without carious lesions," unclear what level of caries they aimed to recruit

Kouchaji 2012 *(Continued)*

	Teeth: permanent first molars Sealants: not reported Surface: occlusal
Patient characteristics and setting	Age: 7 to 12 years, mean 9.5 Sex: 21 male, 19 female Ethnicity: not reported Country: Syria Setting: dental hospital Number of participants/teeth/sites: 40 participants/156 teeth Prevalence: enamel 0.85, dentine 0.29
Index tests	Category of test: DIAGNOdent Sequence of test(s): visual (reference standard) then DIAGNOdent Examiner training and calibration: not reported Teeth cleaning prior to examination: "No prior professional cleaning" Tooth drying prior to examination: 3 to 5 seconds Threshold applied: 0-14 sound, 15-20 enamel, 21+ dentine Device specifics: tip A
Target condition and reference standard(s)	Category: visual Sequence of index test and reference standard: reference standard then index tests Training of examiner: not reported Blinding to index test: not reported Multiple tests: no Site selection: occlusal surface Target condition: Ekstrand criteria
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0
Comparative	
Notes	

Kouchaji 2012 *(Continued)*

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	No		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?			
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	No		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?			
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			

Kouchaji 2012 *(Continued)*

Is the reference standards likely to correctly classify the target condition?	No		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Krause 2007

Study characteristics	
Patient Sampling	Method of sampling: selected "non-cavitated occlusal carious lesions requiring operative intervention (score 3) or teeth where no or preventive treatment was indicated by visual examination and/or bitewing radiographs (scores 0, 1, or 2) were selected" Included conditions: no cavitation or early lesions Teeth: permanent premolars and molars Sealants: no Surface: occlusal
Patient characteristics and setting	Age: mean 36 (+- 8 years) Sex: 34 male, 48 female Ethnicity: not reported Country: Germany Setting: unclear Number of participants/teeth/sites: 82 participants/94 teeth Prevalence: enamel not reported, dentine 0.51
Index tests	Category of test: DIAGNOdent and DIAGNOdent pen Sequence of test(s): visual and radiograph (these determined whether excavation was necessary) then DIAGNOdent/pen

Krause 2007 *(Continued)*

	Examiner training and calibration: not reported Teeth cleaning prior to examination: "professional cleaning of the occlusal surfaces using a rotating soft rubber cup and plain water spray" Tooth drying prior to examination: briefly drying the teeth with air pressure Threshold applied: calculated in study for dentine level only Device specifics: tip specifics not reported
Target condition and reference standard(s)	Category: excavation of those lesions identified through visual and radiograph tests Sequence of index test and reference standard: reference standard then index tests Training of examiner: not reported Blinding to index test: not reported Multiple tests: no Site selection: occlusal surface Target condition: sound, enamel, dentine
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (All)			

Krause 2007 *(Continued)*

Were the index test results interpreted without knowledge of the results of the reference standard?	No		
If a threshold was used, was it pre-specified?	No		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	No		
If a threshold was used, was it pre-specified?	No		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	No		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	No		
Were all patients included in the analysis?	Yes		

Krause 2007 *(Continued)*

Could the patient flow have introduced bias?	High risk

Kucukyilmaz 2015

Study characteristics	
Patient Sampling	Method of sampling: selected Included conditions: non-cavitated and early lesions Teeth: primary molars Sealants: unclear Surface: occlusal
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported Country: Turkey Setting: in vivo study conducted in dental hospital, followed by in vitro after extraction Number of participants/teeth/sites: 200 teeth Prevalence: enamel 0.82, dentine 0.33
Index tests	Category of test: DIAGNOdent Sequence of test(s): visual, radiograph, DIAGNOdent, ECM completed in vivo and in vitro before sectioning of teeth Examiner training and calibration: yes Teeth cleaning prior to examination: "polishes" Tooth drying prior to examination: yes Threshold applied: DIAGNOdent: 0-14 sound and outer enamel, 15-20 inner enamel, 31-30 outer dentine, 31+ deep dentine Device specifics: tip A
Target condition and reference standard(s)	Category: histology Sequence of index test and reference standard: index tests then reference standard Training of examiner: not reported Blinding to index test: not reported Multiple tests: no Site selection: sectioned teeth

Kucukyilmaz 2015 *(Continued)*

	Target condition: sound, outer half of enamel, inner half of enamel, outer half of dentine, deep dentine		
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0		
Comparative			
Notes			
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		

Kucukyilmaz 2015 *(Continued)*

If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Kuhnisch 2006

Study characteristics	
Patient Sampling	Method of sampling: selected Included conditions: non-cavitated and early lesions Teeth: permanent third molars Sealants: no Surface: occlusal
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported Country: Germany

Kuhnisch 2006 *(Continued)*

	Setting: extracted teeth Number of participants/teeth/sites: 54 teeth Prevalence: histology results not clearly reported
Index tests	Category of test: QLF Inspektor Sequence of test(s): index test then reference standard Examiner training and calibration: trained by manufacturer Teeth cleaning prior to examination: brush and polish Tooth drying prior to examination: compressed air Threshold applied: calculated in study using DF, area, and DQ Device specifics: 3 examiners reached agreement in software image
Target condition and reference standard(s)	Category: histology - visual and radiograph of sections Sequence of index test and reference standard: index tests then reference standard Training of examiner: not reported Blinding to index test: same examiner Multiple tests: yes Site selection: sectioned teeth Target condition: sound, outer half of enamel, inner half of enamel, outer half of dentine, deep dentine
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		

Kuhnisch 2006 *(Continued)*

Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	No		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern

Kuhnisch 2006 *(Continued)*

DOMAIN 4: Flow and Timing		
Was there an appropriate interval between index test and reference standard?	Yes	
Did all patients receive the same reference standard?	Yes	
Were all patients included in the analysis?	Yes	
Could the patient flow have introduced bias?		Low risk

Kuhnisch 2007

Study characteristics	
Patient Sampling	Method of sampling: selected, participants already part of ongoing longitudinal study and consented to this additional study after a clinical investigation Included conditions: unclear caries status of participants Teeth: permanent premolars and molars Sealants: yes Surface: occlusal
Patient characteristics and setting	Age: 14 to 15 years Sex: not reported Ethnicity: not reported Country: Erfurt, Germany Setting: school based Number of participants/teeth/sites: 34 participants, 517/311 surfaces/teeth Prevalence: not clearly reported
Index tests	Category of test: QLF Inspektor Sequence of test(s): visual (reference standard) completed prior to index test but examiner independent Examiner training and calibration: "two calibrated investigators" Teeth cleaning prior to examination: unclear Tooth drying prior to examination: 5 seconds air drying Threshold applied: not clearly reported Device specifics: QLF 2.00f software was used to display, score and analyse the images
Target condition and reference standard(s)	Category: visual

Kuhnisch 2007 *(Continued)*

	Sequence of index test and reference standard: index test followed the reference standard Training of examiner: experienced examiners Blinding to index test: yes, clearly stated Multiple tests: no Site selection: occlusal surfaces Target condition: Ekstrand scores: white opacities, brown discolourations, enamel breakdown and dentine exposure
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 206
Comparative	
Notes	Cannot use data, not possible to extract a 2x2 table

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			Unclear
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Unclear		
If multiple tests were applied were different examiners used for each (in vivo)?			
Could the conduct or interpretation of the index test have introduced bias?		Unclear risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern

Kuhnisch 2007 *(Continued)*

DOMAIN 2: Index Test (Green fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Unclear		
If multiple tests were applied were different examiners used for each (in vivo)?			
Could the conduct or interpretation of the index test have introduced bias?		Unclear risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	No		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	No		
Could the patient flow have introduced bias?		High risk	

Kuhnisch 2008

Study characteristics	
Patient Sampling	Method of sampling: selected Included conditions: unclear on exact level of severity to be included in sample Teeth: primary and permanent molars

Kuhnisch 2008 *(Continued)*

	Sealants: yes - labelled where present Surface: occlusal
Patient characteristics and setting	Age: 8 to 12 years Sex: not reported Ethnicity: not reported Country: Germany Setting: primary school Number of participants/teeth/sites: 311 participants/840 occlusal sites Prevalence: 0.71 enamel, 0.06 dentine (ICDAS 4 and above)
Index tests	Category of test: DIAGNOdent Sequence of test(s): reference standard then index test Examiner training and calibration: calibrated Teeth cleaning prior to examination: yes but technique not described Tooth drying prior to examination: 5 seconds air drying Threshold applied: 0-15 sound, 16-17 enamel, 18-31 dentine, 31 deep dentine Device specifics: conical probe A
Target condition and reference standard(s)	Category: visual (ICDAS) Sequence of index test and reference standard: reference standard before index test Training of examiner: calibrated before study Blinding to index test: yes Multiple tests: no Site selection: all occlusal surfaces Target condition: ICDAS
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0
Comparative	
Notes	

Kuhnisch 2008 *(Continued)*

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?			
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Unclear		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?			
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			

Kuhnisch 2008 *(Continued)*

Is the reference standards likely to correctly classify the target condition?	No		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Lee 2018

Study characteristics	
Patient Sampling	Method of sampling: "selected from a pool of extracted human teeth having questionable caries" Included conditions: questionable caries Teeth: permanent premolars and molars Sealants: not reported Surface: occlusal
Patient characteristics and setting	Age: 20 years or older Sex: not reported Ethnicity: not reported Country: South Korea Setting: extracted teeth Number of participants/teeth/sites: 66 teeth (4 were broken during sectioning) Prevalence: 0.81 enamel, 0.11 dentine
Index tests	Category of test: QLF–Digital Biluminator™ 2+, decrease in fluorescence (ΔF) and the increase in red fluorescence (ΔR) are both reported Sequence of test(s): index test (QLF) followed by reference standard Examiner training and calibration: not reported

Lee 2018 *(Continued)*

	Teeth cleaning prior to examination: calculus and soft tissue removed with scaler Tooth drying prior to examination: not reported Threshold applied: optimum thresholds calculated in study: ΔF sound 62, enamel 82, dentine 93 Device specifics: "an analysis patch was delimited by drawing a border that pointed at sound parts without discolorations from the stained pits and fissures with suspected caries according to manufacturer recommendations using the QLF-D software"
Target condition and reference standard(s)	Category: histology Sequence of index test and reference standard: index tests then reference standard Training of examiner: not reported Blinding to index test: not reported Multiple tests: no Site selection: sectioned teeth Target condition: "no enamel demineralization or a narrow surface zone of opacity (scored as 0), enamel demineralization limited to the outer 50% of the enamel layer (scored as 1), demineralization involving the inner 50% of enamel up to the DEJ (scored as 2), and demineralization involving the outer 50% of the dentine (scored as 3)"
Flow and timing	Participants with index test but no reference standard: 4 - reported that these were broken during sectioning Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0
Comparative	
Notes	Data reported for the decrease in fluorescence (ΔF)

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
Could the selection of patients have introduced bias?		High risk	

Lee 2018 *(Continued)*

Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		
If multiple tests were applied were different examiners used for each (in vivo)?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		
If multiple tests were applied were different examiners used for each (in vivo)?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			

Lee 2018 *(Continued)*

Was there an appropriate interval between index test and reference standard?	Yes	
Did all patients receive the same reference standard?	Yes	
Were all patients included in the analysis?	No	
Could the patient flow have introduced bias?		Low risk

Li 2006

Study characteristics	
Patient Sampling	Method of sampling: unclear Included conditions: non-cavitated Teeth: primary molars Sealants: no Surface: occlusal
Patient characteristics and setting	Age: 5 to 6 years, mean 5.3 Sex: not reported Ethnicity: not reported Country: China Setting: school based (kindergarten) Number of participants/teeth/sites: 72 participants/541 teeth Prevalence: not reported
Index tests	Category of test: DIAGNOdent Sequence of test(s): DIAGNOdent then visual Examiner training and calibration: not reported Teeth cleaning prior to examination: yes, with a portable low-speed hand-piece brush Tooth drying prior to examination: yes, dried with high pressure air (triple syringe) Threshold applied: < 10 intact, 10-14 early enamel caries, 15-20 enamel caries, 21-30 early dentine caries, > = 31 deep dentine caries Device specifics: tip A
Target condition and reference standard(s)	Category: visual Sequence of index test and reference standard: following but influenced by DIAGNOdent Training of examiner: not reported

Li 2006 *(Continued)*

	Blinding to index test: no "visual examination was performed on occlusal spots that had the highest laser fluorescence scores during DIAGNOdent examination" Multiple tests: no Site selection: sectioned teeth Target condition: Caries (Ekstrand's index)
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0
Comparative	
Notes	Paper translated by Cochrane author, data not useable as 2x2 table not attainable, study investigates median DIAGNOdent values at each Ekstrand code

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?			
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			

Li 2006 *(Continued)*

DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?			
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	No		
Were the reference standard results interpreted without knowledge of the results of the index tests?	No		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Lussi 1999

Study characteristics	
Patient Sampling	Method of sampling: selected Included conditions: no cavitation, "All teeth had a macroscopically intact occlusal surface" Teeth: not reported Sealants: not reported

Lussi 1999 *(Continued)*

	Surface: occlusal
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported Country: Switzerland Setting: extracted teeth Number of participants/teeth/sites: 105 teeth Prevalence: enamel 0.8, dentine 0.36
Index tests	Category of test: DIAGNOdent Sequence of test(s): index tests (DIAGNOdent then ECM) performed prior to reference standard Examiner training and calibration: not reported Teeth cleaning prior to examination: yes, brush and pumice Tooth drying prior to examination: air dried Threshold applied: calculated in study, 0-4 sound or outer enamel, 5-10 inner enamel,10+ dentine Device specifics: tapered tip
Target condition and reference standard(s)	Category: histology Sequence of index test and reference standard: index test then reference standard Training of examiner: not reported Blinding to index test: unclear Multiple tests: no Site selection: sectioned teeth Target condition: sound, inner/outer enamel, inner/outer dentine
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0
Comparative	
Notes	Results reported at D2 and D3 thresholds so not relevant to our primary outcome
Methodological quality	

Lussi 1999 *(Continued)*

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		
If multiple tests were applied were different examiners used for each (in vivo)?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		
If multiple tests were applied were different examiners used for each (in vivo)?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		

Lussi 1999 *(Continued)*

Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Lussi 2001

Study characteristics	
Patient Sampling	Method of sampling: dentists selected participants but method or criteria unclear Included conditions: aims of inclusion not clearly stated Teeth: not reported Sealants: not reported Surface: occlusal
Patient characteristics and setting	Age: mean age 19.8 years Sex: not reported Ethnicity: not reported Country: Switzerland and Germany Setting: clinical setting Number of participants/teeth/sites: 240 participants/332 surfaces Prevalence: enamel 0.67, dentine 0.59
Index tests	Category of test: DIAGNOdent Sequence of test(s): visual, radiograph then DIAGNOdent Examiner training and calibration: experienced examiners, with training Teeth cleaning prior to examination: "Professional cleaning of the tooth surfaces was not carried out. If needed, plaque remnants were removed from the fissures using an explorer"

Lussi 2001 *(Continued)*

	Tooth drying prior to examination: air dried Threshold applied: calculated in study, 0-13: no caries; values 14-20: enamel caries; values > 20: dentinal caries Device specifics: tip A
Target condition and reference standard(s)	Category: excavation for those deemed to be appropriate, not clearly reported how this decision was made, it appears that visual, radiograph, and DIAGNOdent were combined to inform this decision Sequence of index test and reference standard: visual, radiograph then DIAGNOdent, all before excavation Training of examiner: experienced clinicians Blinding to index test: no - reference standard appears to be directly informed by the index test Multiple tests: yes Site selection: via clinical decision making and combined series of tests Target condition: sound, enamel, inner/outer dentine
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	No		

Lussi 2001 *(Continued)*

If a threshold was used, was it pre-specified?	No		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	No		
If a threshold was used, was it pre-specified?	No		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	No		
Were the reference standard results interpreted without knowledge of the results of the index tests?	No		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	No		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		High risk	

Lussi 2003

Study characteristics	
Patient Sampling	Method of sampling: selected Included conditions: no cavitation "macroscopically intact occlusal surface" Teeth: not reported, "extracted deciduous teeth" Sealants: not reported Surface: occlusal
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported Country: Switzerland Setting: extracted teeth Number of participants/teeth/sites: 95 teeth Prevalence: enamel 0.85, dentine 0.18
Index tests	Category of test: DIAGNOdent Sequence of test(s): index tests (visual, visual (telescope), visual (probe), bitewing radiograph, DIAGNOdent) performed prior to reference standard Examiner training and calibration: 3 experienced dentists Teeth cleaning prior to examination: yes, water and brush Tooth drying prior to examination: air dried for 2 seconds Threshold applied: calculated in study, 0-4 sound or outer enamel, 5-12 inner enamel, 12+ dentine Device specifics: tapered tip
Target condition and reference standard(s)	Category: histology Sequence of index test and reference standard: index test then reference standard Training of examiner: not reported Blinding to index test: unclear Multiple tests: no Site selection: labelled drawing and sectioned teeth Target condition: sound, inner/outer enamel, inner/outer dentine
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0

Lussi 2003 *(Continued)*

	Time interval between tests: minimal Participants receiving both tests but excluded from results: 0		
Comparative			
Notes			
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		
If multiple tests were applied were different examiners used for each (in vivo)?	No		

Lussi 2003 *(Continued)*

Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Lussi 2005

Study characteristics	
Patient Sampling	Method of sampling: dentists selected participants but method unclear Included conditions: "occlusal macroscopically intact surfaces" Teeth: premolars and molars Sealants: not reported Surface: occlusal
Patient characteristics and setting	Age: mean age 18 years Sex: not reported Ethnicity: not reported Country: Switzerland Setting: clinical setting Number of participants/teeth/sites: 70 participants/117 surfaces

Lussi 2005 *(Continued)*

	Prevalence: enamel 0.68, dentine 0.64
Index tests	Category of test: DIAGNOdent Sequence of test(s): index test prior to reference standard Examiner training and calibration: experienced examiners, with training Teeth cleaning prior to examination: study designed to assess levels of cleaning: "(1) moist, uncleaned surface; (2) dried, uncleaned surface; (3) moist, cleaned surface; (4) dried, cleaned surface. PROPHYflex 2 (KaVo) with NaHCO 3 powder and water was used to clean the occlusal surface for 5 s" Tooth drying prior to examination: air dried as to allow for dry and moist measurements Threshold applied: calculated in study, 0-16: no caries; values 16-18: enamel caries; values > 32: dentinal caries (taken at the clean and dried results) Device specifics: tip A
Target condition and reference standard(s)	Category: excavation for those deemed to be appropriate, not clearly reported how this decision was made, it appears that visual and existing radiographs were combined to inform this decision Sequence of index test and reference standard: visual, radiograph then DIAGNOdent, all before excavation Training of examiner: experienced clinicians Blinding to index test: no - reference standard appears to be directly informed by the index test Multiple tests: yes Site selection: via clinical decision making and combined series of tests Target condition: sound, enamel, inner/outer dentine
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			

Lussi 2005 *(Continued)*

Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	No		
If a threshold was used, was it pre-specified?	No		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	No		
If a threshold was used, was it pre-specified?	No		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	No		
Were the reference standard results interpreted without knowledge of the results of the index tests?	No		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	

Lussi 2005 *(Continued)*

Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	No		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		High risk	

Lussi 2006

Study characteristics	
Patient Sampling	Method of sampling: selected Included conditions: no cavitation Teeth: permanent molars Sealants: not reported Surface: approximal
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported Country: Switzerland Setting: extracted teeth Number of participants/teeth/sites: 75 teeth/150 sites Prevalence: enamel 0.59, dentine 0.25
Index tests	Category of test: DIAGNOdent pen, "The roots were embedded in composite to arrange these three teeth in a manner that simulated contact points of adult teeth" Sequence of test(s): index tests (bitewing radiograph then DIAGNOdent pen) performed prior to reference standard Examiner training and calibration: not reported Teeth cleaning prior to examination: water and brush 15 seconds, 10 seconds prophylex and sodium bicarbonate Tooth drying prior to examination: not reported Threshold applied: calculated in study, 2 tips investigated: wedge: 0-6 sound, 6.1-9 outer enamel, 9.1-15 inner enamel, 15+ dentine

Lussi 2006 *(Continued)*

	tapered: 0-9 sound, 9.1-13 outer enamel, 13.1-19 inner enamel, 22+ dentine Device specifics: 2 sapphire tips 0.4 mm, and 1.1 mm and 0.7 mm (tapered)
Target condition and reference standard(s)	Category: histology Sequence of index test and reference standard: index test then reference standard Training of examiner: not reported Blinding to index test: unclear Multiple tests: no Site selection: sectioned teeth Target condition: sound, inner/outer enamel, inner/outer dentine
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0
Comparative	
Notes	Data extracted for wedge tip

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		
If multiple tests were applied were different examiners used for each (in vivo)?	No		

Lussi 2006 *(Continued)*

Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Lussi 2006a

Study characteristics

Lussi 2006a *(Continued)*

Patient Sampling	Method of sampling: selected Included conditions: no cavitation and early cavitation Teeth: permanent third molars Sealants: not reported Surface: occlusal
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported Country: Switzerland Setting: extracted teeth Number of participants/teeth/sites: 119 sites Prevalence: enamel 0.78, dentine 0.35
Index tests	Category of test: DIAGNOdent and DIAGNOdent pen Sequence of test(s): index tests (DIAGNOdent and DIAGNOdent pen - cylindrical and conical tips) performed prior to reference standard Examiner training and calibration: not reported Teeth cleaning prior to examination: water and brush 15 seconds, 10 seconds prophylex and sodium bicarbonate Tooth drying prior to examination: not reported Threshold applied: calculated within study: DIAGNOdent: 0-7 sound, 7.1-14 outer enamel, 14.1-24 inner enamel, 24+ dentine DIAGNOdentpen: cylindrical tip: 0-6 sound, 6.1-13 outer enamel, 13.1-17 inner enamel, 17+ dentine; conical tip: 0-7 sound, 7.1-12 outer enamel, 12.1-19 inner enamel, 19+ dentine Device specifics: DIAGNOdent: tip A DIAGNOdentpen: 2 tips: cylindrical and conical tips
Target condition and reference standard(s)	Category: histology Sequence of index test and reference standard: index test then reference standard Training of examiner: not reported Blinding to index test: unclear Multiple tests: no Site selection: sectioned teeth

Lussi 2006a *(Continued)*

	Target condition: sound, inner/outer enamel, inner/outer dentine
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		

Lussi 2006a *(Continued)*

If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Mansour 2016

Study characteristics	
Patient Sampling	Method of sampling: unclear Included conditions: no cavitation and early cavitation, "Subjects with open cavities extending into dentin were excluded" Teeth: permanent third molars Sealants: not reported Surface: "all coronal areas of the teeth considered to be at high risk of caries: occlusal and approximal, white or brown spot lesions, non-cavitated and cavitated potential lesions, fissures, and adjacent to restorations"
Patient characteristics and setting	Age: 19 to 52 years, mean 34 Sex: 16 male, 24 female Ethnicity: not reported

Mansour 2016 *(Continued)*

	Country: US Setting: dental clinic Number of participants/teeth/sites: 40 participants/932 teeth (426 untreated teeth used in this sample) Prevalence: untreated teeth: enamel 0.12; previously treated: enamel 0.14
Index tests	Category of test: DIAGNOdent Sequence of test(s): index tests (OCT also completed and potentially interpreted by same examiner) performed after reference standard Examiner training and calibration: 90-minute training session Teeth cleaning prior to examination: not reported Tooth drying prior to examination: not reported Threshold applied: diagnostic limits were set at the levels prescribed by the manufacturer
Target condition and reference standard(s)	Category: visual and radiograph "detailed dental examination by one experienced clinician using loupes (2.5 magnification), and radiographs according to standard clinical practice" Sequence of index test and reference standard: index test followed reference standard Training of examiner: not reported Blinding to index test: unclear Multiple tests: yes Site selection: sectioned teeth Target condition: "Teeth were considered carious if there were white or brown spot lesions on the tooth not consistent with the clinical appearance of sound enamel" “healthy” being scored if both observers scored healthy, and “not-healthy” scored if one or both observers scored “not-healthy”
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0
Comparative	
Notes	Untreated teeth used in the data extraction for analysis

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			

Mansour 2016 *(Continued)*

Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	No		
If a threshold was used, was it pre-specified?	Unclear		
If multiple tests were applied were different examiners used for each (in vivo)?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		Unclear risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Unclear
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	No		
If a threshold was used, was it pre-specified?	Unclear		
If multiple tests were applied were different examiners used for each (in vivo)?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		Unclear risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Unclear
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	No		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		

Mansour 2016 *(Continued)*

Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Manton 2007

Study characteristics	
Patient Sampling	Method of sampling: selected Included conditions: unclear; "free of enamel defects or evidence of gross caries" Teeth: permanent, third molars Sealants: yes Surface: occlusal
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported Country: Australia Setting: extracted teeth Number of participants/teeth/sites: 67 teeth Prevalence: enamel 0.68, dentine 0.23
Index tests	Category of test: DIAGNOdent Sequence of test(s): visual, radiograph, FOTI, LF, tactile Examiner training and calibration: none Teeth cleaning prior to examination: soft tissue removed Tooth drying prior to examination: unclear Threshold applied: 0-13 sound, 14-20 enamel, > 20 dentine

Manton 2007 *(Continued)*

	Device specifics: tip A		
Target condition and reference standard(s)	Category: histology Sequence of index test and reference standard: index test then reference standard Training of examiner: not reported Blinding to index test: not reported Multiple tests: no Site selection: sectioned teeth Target condition: sound, outer half of enamel, inner half of enamel, to dentino-enamel junction, halfway between dentino-enamel junction and pulp, greater than halfway between dentino-enamel junction and pulp		
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0		
Comparative			
Notes	Sensitivity and specificity presented data at dentine level only		
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		

Manton 2007 *(Continued)*

If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Markowitz 2013

Study characteristics	
Patient Sampling	Method of sampling: selected Included conditions: "Teeth needed to have areas of the pits and fissures lesions classified as ICDAS code 2 or 3 (having distinct colour change present when wet and possible local enamel breakdown but without shadow in the underlying dentine)" Teeth: permanent, third molars Sealants: yes Surface: occlusal
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported Country: USA Setting: extracted teeth Number of participants/teeth/sites: 31 teeth Prevalence: not reported
Index tests	Category of test: Spectra™ Caries Detection Aid, a fluorescent camera Sequence of test(s): index test then reference standard Examiner training and calibration: yes Teeth cleaning prior to examination: fine pumice and brush Tooth drying prior to examination: damp surfaces Threshold applied: 0.0 to 0.9 - green - sound 1.0 to 1.4 - blue - initial enamel lesions 1.5 to 1.9 - red - enamel lesions up to EDJ 2.0 to 2.4 - orange - dentine lesions > 2.5 - yellow - deep dentine lesions Device specifics: image of entire surface, mean of peak fluorescent camera reading, 10 mm spacer and infection control sleeve
Target condition and reference standard(s)	Category: histology Sequence of index test and reference standard: following index test Training of examiner: not reported Blinding to index test: not reported

Markowitz 2013 *(Continued)*

	Multiple tests: no Site selection: unclear Target condition: sound, enamel, dentine
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	No		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?			
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		

Markowitz 2013 *(Continued)*

If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?			
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Red fluorescence)			
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Unclear		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Unclear
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Markowitz 2015

Study characteristics	
Patient Sampling	Method of sampling: selected Included conditions: "teeth lacking visually apparent caries and teeth with small lesions" Teeth: permanent, third molars Sealants: yes Surface: occlusal
Patient characteristics and setting	Age: not reported Sex: not reported

Markowitz 2015 *(Continued)*

	Ethnicity: not reported Country: USA Setting: extracted teeth Number of participants/teeth/sites: 90 teeth/sites Prevalence: enamel 1.00, dentine unclear
Index tests	Category of test: Spectra™ Caries Detection Aid, a fluorescent camera Sequence of test(s): index test then reference standard Examiner training and calibration: yes Teeth cleaning prior to examination: fine pumice and brush Tooth drying prior to examination: not reported Threshold applied: 0.0-0.9 - green - sound 1.0-1.4 - blue - initial enamel lesions 1.5-1.9 - red - enamel lesions up to EDJ 2.0-2.4 - orange - dentine lesions > 2.5 - yellow - deep dentine lesions Device specifics: "using uniform examination methods and positioning"
Target condition and reference standard(s)	Category: histology Sequence of index test and reference standard: following index test Training of examiner: not reported Blinding to index test: not reported Multiple tests: no Site selection: cut along EDJ, different method to all other studies Target condition: dentine caries
Flow and timing	Participants with index test but no reference standard: not reported Participants with reference standard but no index test: not reported Time interval between tests: not reported Participants receiving both tests but excluded from results: not reported
Comparative	

Markowitz 2015 *(Continued)*

Notes

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?			
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?			
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Red fluorescence)			
DOMAIN 3: Reference Standard			

Markowitz 2015 *(Continued)*

Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			High
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Unclear		
Could the patient flow have introduced bias?		Unclear risk	

Matos 2011

Study characteristics	
Patient Sampling	Method of sampling: randomly selected from available patients Included conditions: "children seeking dental treatment at the School of Dentistry of the University of São Paulo were selected" Teeth: primary, molars Sealants: not reported Surface: occlusal
Patient characteristics and setting	Age: 4 to 12 years Sex: 30 male, 38 female Ethnicity: not reported Country: Brazil Setting: dental hospital patients Number of participants/teeth/sites: 383 teeth in 68 participants Prevalence: enamel 0.91, dentine 0.05
Index tests	Category of test: DIAGNOdent pen and VistaProof Sequence of test(s): visual inspection and radiographic methods, then LF pen, fluorescence camera Examiner training and calibration: yes

Matos 2011 *(Continued)*

	Teeth cleaning prior to examination: rotating bristle brush and a pumice/water slurry Tooth drying prior to examination: standardized drying time of 5 seconds Threshold applied: DIAGNOdent pen: sound 0–4; enamel lesions > 4; dentine lesion > 34 VistaProof: sound 0–1.1; enamel lesions > 1.1; dentine lesions > 1.4 Device specifics: DIAGNOdent pen: probe tip 2 VitsaProof: "the image of each surface was recorded by the camera software (DBSWIN, Dürr Dental)"
Target condition and reference standard(s)	Category: visual for enamel threshold, excavation and visual used for dentine threshold Sequence of index test and reference standard: following index test, but performed by same examiner during the same appointment Training of examiner: yes Blinding to index test: unclear - the same examiner performed all tests so difficult to blind results Multiple tests: yes Site selection: a drawing of the occlusal surface was made to indicate the selected site Target condition: sound, enamel, dentine
Flow and timing	Participants with index test but no reference standard: 25 for enamel caries as examiners did not agree Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		

Matos 2011 *(Continued)*

Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	No		
If a threshold was used, was it pre-specified?	No		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	No		
If a threshold was used, was it pre-specified?	No		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Red fluorescence)			
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	No		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			

Matos 2011 *(Continued)*

Was there an appropriate interval between index test and reference standard?	Yes	
Did all patients receive the same reference standard?	No	
Were all patients included in the analysis?	No	
Could the patient flow have introduced bias?		High risk

Mendes 2005

Study characteristics	
Patient Sampling	Method of sampling: selected Included conditions: non-cavitated; "surfaces with no clinical signs of caries or with white spot caries lesions" Teeth: primary molars Sealants: unclear Surface: smooth
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported Country: Brazil Setting: extracted teeth Number of participants/teeth/sites: 77 sites Prevalence: enamel 0.86, dentine 0.14
Index tests	Category of test: DIAGNOdent Sequence of test(s): index test then reference standard Examiner training and calibration: not reported Teeth cleaning prior to examination: toothbrush and water Tooth drying prior to examination: not reported Threshold applied: calculated within study: 0-3 sound, 4-7 enamel, > 8 dentine Device specifics: tip B
Target condition and reference standard(s)	Category: histology Sequence of index test and reference standard: index test then reference standard Training of examiner: not reported

Mendes 2005 *(Continued)*

	Blinding to index test: not reported Multiple tests: no Site selection: marked on tooth then sectioned Target condition: sound, outer half of enamel, inner half of enamel, outer half of dentine, deep dentine		
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0		
Comparative			
Notes			
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			

Mendes 2005 *(Continued)*

DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Mendes 2006

Study characteristics	
Patient Sampling	Method of sampling: selected Included conditions: non-cavitated Teeth: primary molars Sealants: unclear Surface: occlusal
Patient characteristics and setting	Age: not reported

Mendes 2006 *(Continued)*

	Sex: not reported Ethnicity: not reported Country: Brazil Setting: extracted teeth Number of participants/teeth/sites: 79 teeth/110 sites Prevalence: enamel 0.75, dentine 0.25
Index tests	Category of test: DIAGNOdent Sequence of test(s): visual, then DIAGNOdent, then radiograph Examiner training and calibration: not reported Teeth cleaning prior to examination: brush, pumice and slurry Tooth drying prior to examination: 5 seconds Threshold applied: calculated within study: 0-7 sound, 8-14 enamel, > 14 dentine Device specifics: tip A, sites marked on photograph, then maximum value of read-out taken, mean of 3 measurements for final result
Target condition and reference standard(s)	Category: histology Sequence of index test and reference standard: index test then reference standard Training of examiner: not reported Blinding to index test: not reported Multiple tests: no Site selection: marked on tooth then sectioned Target condition: sound, outer half of enamel, inner half of enamel, outer half of dentine, deep dentine
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0
Comparative	
Notes	
Methodological quality	

Mendes 2006 *(Continued)*

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		
If multiple tests were applied were different examiners used for each (in vivo)?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		
If multiple tests were applied were different examiners used for each (in vivo)?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		

Mendes 2006 *(Continued)*

Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Mendes 2012

Study characteristics	
Patient Sampling	Method of sampling: selected Included conditions: non-cavitated and enamel lesions Teeth: primary molars Sealants: unclear Surface: occlusal and approximal
Patient characteristics and setting	Age: 4 to 12 years, mean 7.3 Sex: occlusal - 30 male, 38 female; approximal - 53 male, 73 female Ethnicity: not reported Country: Brazil Setting: dental hospital Number of participants/teeth/sites: occlusal - 68 participants/407 sites; proximal - 132 participants/1213 sites Prevalence: occlusal - dentine 0.05; proximal - 0.04
Index tests	Category of test: DIAGNOdent Sequence of test(s): visual, then radiograph, then DIAGNOdent Examiner training and calibration: not reported Teeth cleaning prior to examination: brush, pumice and slurry Tooth drying prior to examination: 5 seconds Threshold applied: occlusal: > 34 dentine; approximal: > 16

Mendes 2012 *(Continued)*

	Device specifics: tip 2 was used for occlusal surfaces; tip 1 was used for approximal surfaces
Target condition and reference standard(s)	Category: occlusal - excavation of those teeth suspected of dentinal caries (ICDAS score of 6 3), the remainder received visual assessment only approximal - "temporary separation using orthodontic rubber rings placed around the contact points for 7 days. Two examiners evaluated each surface for the presence of cavities" Sequence of index test and reference standard: index test then reference standard Training of examiner: not reported Blinding to index test: not reported Multiple tests: yes Site selection: photographed and site selected Target condition: cavitated caries lesions; "the cut-off point for visual inspection was an ICDAS score of 3"
Flow and timing	Participants with index test but no reference standard: occlusal: unclear, approximal: 6 Participants with reference standard but no index test: 0 Time interval between tests: occlusal - unclear, approximal - 1 week Participants receiving both tests but excluded from results: 0
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	No		

Mendes 2012 *(Continued)*

If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	No		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	No		
Were the reference standard results interpreted without knowledge of the results of the index tests?	No		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	No		
Were all patients included in the analysis?	No		
Could the patient flow have introduced bias?		High risk	

Mepparambath 2014

Study characteristics	
Patient Sampling	Method of sampling: selected Included conditions: non-cavitated and enamel lesions, "primary molars without obvious cavities were identified in children" Teeth: primary molars Sealants: unclear Surface: approximal
Patient characteristics and setting	Age: 3 to 10 years Sex: not reported Ethnicity: not reported Country: India Setting: dental school Number of participants/teeth/sites: 101 teeth/169 sites Prevalence: enamel 0.22, dentine 0.08
Index tests	Category of test: DIAGNOdent Sequence of test(s): index test then reference standard Examiner training and calibration: not reported Teeth cleaning prior to examination: rubber cup with pumice Tooth drying prior to examination: 3-way syringe Threshold applied: calculated within study: 0-9 sound, 10-17 enamel, 18+ dentine Device specifics: probe A
Target condition and reference standard(s)	Category: bitewing radiograph Sequence of index test and reference standard: index test then reference standard Training of examiner: not reported Blinding to index test: not reported Multiple tests: no Site selection: marked on tooth then sectioned Target condition: sound, outer enamel, inner enamel, past EDJ
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0

Mepparambath 2014 *(Continued)*

	Time interval between tests: minimal Participants receiving both tests but excluded from results: 0		
Comparative			
Notes			
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?			
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?			

Mepparambath 2014 *(Continued)*

Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	No		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Mortensen 2018

Study characteristics	
Patient Sampling	Method of sampling: sites selected in each participant Included conditions: non-cavitated and enamel lesions, "various stages of occlusal caries" Teeth: permanent molars Sealants: unclear Surface: occlusal
Patient characteristics and setting	Age: 20 to 66 years Sex: 21% male Ethnicity: not reported Country: Denmark Setting: university setting: patients, employees, and students Number of participants/teeth/sites: 57 participants/60 sites

Mortensen 2018 *(Continued)*

	Prevalence: enamel 0.97, dentine 0.45
Index tests	Category of test: DIAGNOdent pen Sequence of test(s): index test then reference standard, ordered: ECM (CarieScan), then DIAGNOdent pen then visual and radiograph Examiner training and calibration: experienced and trained Teeth cleaning prior to examination: rotating brush Tooth drying prior to examination: 5 seconds Threshold applied: calculated within study: 0-12 sound, 13-24 enamel, 25+ dentine Device specifics: cylindrical probe
Target condition and reference standard(s)	Category: visual (ICDAS) Sequence of index test and reference standard: reference standard follows ECM and DIAGNOdent pen Training of examiner: experienced examiners Blinding to index test: no Multiple tests: no, only visual used Site selection: first examiner labelled the location on a plan Target condition: ICDAS 1 to 5
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
Could the selection of patients have introduced bias?		High risk	

Mortensen 2018 *(Continued)*

Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	No		
Were the reference standard results interpreted without knowledge of the results of the index tests?	No		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		

Mortensen 2018 *(Continued)*

Did all patients receive the same reference standard?	Yes	
Were all patients included in the analysis?	Yes	
Could the patient flow have introduced bias?		Low risk

Muller-Bolla 2017

Study characteristics	
Patient Sampling	Method of sampling: unclear how participants were identified Included conditions: enamel lesions, possibly cavitated, "Caries-free subjects (without carious lesions diagnosed by both visual examination and bitewing radiographs) or uncooperative children during the examination were excluded" Teeth: primary and permanent, premolars and molars Sealants: unclear Surface: occlusal
Patient characteristics and setting	Age: 5 to 15 years Sex: 60% male Ethnicity: not reported Country: France Setting: university hospital, attending paediatric clinic Number of participants/teeth/sites: 103 participants/743 sites Prevalence: enamel 0.72, dentine 0.29
Index tests	Category of test: DIAGNOdent pen and Soprolife Sequence of test(s): visual and radiograph (reference standard) followed by Soprolife and DIAGNOdent pen Examiner training and calibration: 1 day calibration session Teeth cleaning prior to examination: sodium bicarbonate powder-cleaning tool was used for 5 to 10 seconds per tooth Tooth drying prior to examination: not reported Threshold applied: Soprolife: sound - shiny green, outer enamel - tiny, thin red or grey shimmer in the pits and fissure, inner enamel - red shimmer, grey or black colouration in the pits and fissure, dentine - red areas wider than fissures; surface roughness occurs, possibly grey or rough grey zone visible DIAGNOdent pen: 0-12 sound, 13-24 enamel, 25+ dentine Device specifics: fibre tip for DIAGNOdent pen, Soprolife - "studied using the SoproImaging software"

Muller-Bolla 2017 *(Continued)*

Target condition and reference standard(s)	Category: visual (ICDAS) and radiograph Sequence of index test and reference standard: reference standard before Soprolife and DIAGNOdent pen Training of examiner: experienced examiners with 1 day training Blinding to index test: yes Multiple tests: yes Site selection: full surface assessed Target condition: ICDAS 1 to 6
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0
Comparative	
Notes	Prevalence of caries for primary and permanent dentition unknown so have to use the mixed dentition results

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	No		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	No		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	No		

Muller-Bolla 2017 *(Continued)*

Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	No		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Red fluorescence)			
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	No		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Neuhaus 2011

Study characteristics

Neuhaus 2011 *(Continued)*

Patient Sampling	Method of sampling: selected Included conditions: unclear Teeth: primary molars (first and second) Sealants: unclear Surface: occlusal
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported Country: Switzerland Setting: extracted teeth Number of participants/teeth/sites: 37 teeth/37 sites Prevalence: enamel 0.73, dentine 0.24
Index tests	Category of test: DIAGNOdent and DIAGNOdent pen Sequence of test(s): index tests (visual, DIAGNOdent devices then radiograph) then reference standard Examiner training and calibration: experienced examiners Teeth cleaning prior to examination: 3 in 1 syringe Tooth drying prior to examination: not reported Threshold applied: "D1 and D3 were determined according to se and sp" DIAGNOdent: 0-9 sound, 10-11 enamel, 17+ dentine DIAGNOdent pen: 0-13 sound, 14-30 enamel, 31+ dentine Device specifics: DIAGNOdent - tip A; DIAGNOdent pen - cylindrical sapphire fibre tip
Target condition and reference standard(s)	Category: histology Sequence of index test and reference standard: index test then reference standard Training of examiner: not reported Blinding to index test: not reported Multiple tests: no Site selection: unclear Target condition: sound, outer half of enamel, inner half of enamel, outer half of dentine, deep dentine
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0

Neuhaus 2011 *(Continued)*

	Time interval between tests: minimal Participants receiving both tests but excluded from results: 0		
Comparative			
Notes			
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		
If multiple tests were applied were different examiners used for each (in vivo)?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		
If multiple tests were applied were different examiners used for each (in vivo)?	Unclear		

Neuhaus 2011 *(Continued)*

Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Novaes 2009

Study characteristics	
Patient Sampling	Method of sampling: random Included conditions: no cavitation and early lesions Teeth: primary molars (first and second) Sealants: unclear Surface: approximal
Patient characteristics and setting	Age: 5 to 12 years, mean 7.7 Sex: 21 male, 29 female Ethnicity: not reported Country: Brazil Setting: dental hospital Number of participants/teeth/sites: 50 participants/621 sites

Novaes 2009 *(Continued)*

	Prevalence: enamel 0.41, dentine 0.03
Index tests	Category of test: DIAGNOdent pen Sequence of test(s): index tests (visual, radiograph, DIAGNOdent pen) then reference standard Examiner training and calibration: trained but no calibration Teeth cleaning prior to examination: brush and slurry Tooth drying prior to examination: air dried, 5 seconds Threshold applied: calculated within study; 0–5 sound; 5.1–16 white-spot caries; 16+ cavitation Device specifics: tip A
Target condition and reference standard(s)	Category: visual - separators Sequence of index test and reference standard: index test then reference standard Training of examiner: not reported Blinding to index test: unclear Multiple tests: no Site selection: approximal surface Target condition: sound, white spot, cavitated
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: 1 week to allow for separation of teeth Participants receiving both tests but excluded from results: 0
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	

Novaes 2009 *(Continued)*

Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	No		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		

Novaes 2009 *(Continued)*

Did all patients receive the same reference standard?	Yes	
Were all patients included in the analysis?	Yes	
Could the patient flow have introduced bias?		Low risk

Novaes 2010

Study characteristics	
Patient Sampling	Method of sampling: randomly selected, although precise methods are unclear Included conditions: no cavitation and early lesions Teeth: primary molars Sealants: unclear Surface: approximal, "exams were performed on the distal face of first primary molars, the mesial face of second primary molars and also the distal face of second primary molars"
Patient characteristics and setting	Age: 4 to 12 years, mean 7.25 Sex: 32 male, 44 female Ethnicity: not reported Country: Sao Paulo, Brazil Setting: dental hospital Number of participants/teeth/sites: 76 participants/168 teeth/592 sites Prevalence: enamel 0.81, dentine 0.05
Index tests	Category of test: DIAGNOdent pen Sequence of test(s): index tests (randomly ordered: visual, radiograph, DIAGNOdent pen) then reference standard Examiner training and calibration: trained but no calibration Teeth cleaning prior to examination: brush and slurry Tooth drying prior to examination: air dried, 5 seconds Threshold applied: calculated within study; 0–5 sound; 5.1–16 white-spot caries; 16+ cavitation Device specifics: tip 1
Target condition and reference standard(s)	Category: visual - separators Sequence of index test and reference standard: index tests then reference standard Training of examiner: not reported

Novaes 2010 *(Continued)*

	Blinding to index test: unclear Multiple tests: no Site selection: approximal surface Target condition: sound, white spot, cavitated
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: 1 week to allow for separation of teeth Participants receiving both tests but excluded from results: 0
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			

Novaes 2010 *(Continued)*

DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	No		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Novaes 2012

Study characteristics	
Patient Sampling	Method of sampling: randomly selected, although precise methods are unclear, "randomly selected using the enrolment or history form of each child" Included conditions: no cavitation and early lesions Teeth: primary molars - first and second present Sealants: unclear Surface: approximal

Novaes 2012 *(Continued)*

Patient characteristics and setting	Age: 4 to 12 years, mean 7.4 Sex: 32 male, 44 female Ethnicity: not reported Country: Sao Paulo, Brazil Setting: dental hospital, "children seeking dental treatment" suggests there will be some caries Number of participants/teeth/sites: 76 participants/344 approximal "spaces"/520 surfaces Prevalence: enamel 0.8
Index tests	Category of test: DIAGNOdent pen Sequence of test(s): index tests (visual, radiograph, DIAGNOdent pen) then reference standard Examiner training and calibration: trained but no calibration Teeth cleaning prior to examination: not reported Tooth drying prior to examination: not reported Threshold applied: calculated within study; 0–5 sound; 6+ cavitation Device specifics: tip A
Target condition and reference standard(s)	Category: visual - separators Sequence of index test and reference standard: index tests then reference standard Training of examiner: yes Blinding to index test: unclear Multiple tests: no Site selection: approximal surface Target condition: sound and caries (including: white spot, cavitation)
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: 1 week to allow for separation of teeth Participants receiving both tests but excluded from results: 0
Comparative	
Notes	Not possible to use data, 2x2 table not possible to construct since the outcome of interest was the affect of spacing on index test
Methodological quality	

Novaes 2012 *(Continued)*

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		Unclear risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		Unclear risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	No		

Novaes 2012 *(Continued)*

Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Novaes 2012a

Study characteristics	
Patient Sampling	Method of sampling: selected Included conditions: unclear Teeth: primary molars - "recently extracted primary molars were selected" Sealants: unclear Surface: occlusal
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported Country: Sao Paulo, Brazil Setting: extracted teeth Number of participants/teeth/sites: 77 teeth/113 sites Prevalence: enamel 0.57, dentine 0.17
Index tests	Category of test: DIAGNOdent, DIAGNOdentpen and VistaProof Sequence of test(s): index tests (radiograph,visual, DIAGNOdent, VistaProof) then reference standard Examiner training and calibration: trained but no calibration Teeth cleaning prior to examination: yes Tooth drying prior to examination: yes

Novaes 2012a *(Continued)*

	Threshold applied: calculated in study: DIAGNOdent: 0–7 sound, 8-23 enamel, 24+ dentine DIAGNOdentpen: 0-8 sound, 9-30 enamel, 31+ dentine VistaProof: "numerical value from 0 to 3 corresponding to the lesion severity is assigned" Device specifics: tip A
Target condition and reference standard(s)	Category: histology Sequence of index test and reference standard: index test then reference standard Training of examiner: not reported Blinding to index test: yes Multiple tests: no Site selection: sectioned teeth Target condition: sound, outer half of enamel, inner half of enamel, outer half of dentine, deep dentine
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: 1 week to allow for separation of teeth Participants receiving both tests but excluded from results: 0
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All)			

Novaes 2012a *(Continued)*

Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		
If multiple tests were applied were different examiners used for each (in vivo)?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		
If multiple tests were applied were different examiners used for each (in vivo)?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Red fluorescence)			
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		

Novaes 2012a *(Continued)*

Could the patient flow have introduced bias?	Low risk

Novaes 2016

Study characteristics	
Patient Sampling	Method of sampling: selected Included conditions: non-cavitated and enamel lesions Teeth: primary molars - "recently extracted or exfoliated primary molars were selected" Sealants: unclear Surface: approximal
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported Country: Brazil Setting: extracted teeth Number of participants/teeth/sites: 65 teeth/99 sites Prevalence: enamel 0.7, dentine 0.23
Index tests	Category of test: DIAGNOdent, DIAGNOdent pen and VistaProof Sequence of test(s): index tests (VistaProof, DIAGNOdent, DIAGNOdent pen; 1 week apart) then reference standard Examiner training and calibration: experienced Teeth cleaning prior to examination: brush and slurry Tooth drying prior to examination: air dry 3 seconds Threshold applied: "receiver operating characteristic curve (ROC) analysis was used to determine the best cutoff points for the devices at each threshold (D1, D2, and D3)": DIAGNOdent: 0–2 sound; 3-21 enamel; 22+ dentine DIAGNOdent pen: 0-3 sound; 4-19 enamel; 20+ dentine Vista Proof: 0-1.1 sound; 1.2-1.6 enamel; 1.7+ dentine Device specifics: DIAGNOdent: tip B DIAGNOdent pen: tip 1
Target condition and reference standard(s)	Category: histology Sequence of index test and reference standard: index test then reference standard

Novaes 2016 *(Continued)*

	Training of examiner: not reported Blinding to index test: yes Multiple tests: no Site selection: sectioned teeth Target condition: sound, enamel, dentine		
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: 1 week to allow for separation of teeth Participants receiving both tests but excluded from results: 0		
Comparative			
Notes	Results taken for examiner 1		
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		
If multiple tests were applied were different examiners used for each (in vivo)?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			

Novaes 2016 *(Continued)*

DOMAIN 2: Index Test (Blue fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		
If multiple tests were applied were different examiners used for each (in vivo)?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Red fluorescence)			
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Ouellet 2002

Study characteristics	
Patient Sampling	Method of sampling: selected Included conditions: non-cavitated and enamel lesions "questionable occlusal carious lesions" Teeth: permanent third molars Sealants: unclear

Ouellet 2002 *(Continued)*

	Surface: occlusal
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported Country: Canada Setting: extracted teeth Number of participants/teeth/sites: 100 teeth Prevalence: not clearly reported
Index tests	Category of test: DIAGNOdent Sequence of test(s): index tests (DIAGNOdent pen followed by visual examination) followed by reference standard Examiner training and calibration: 1 examiner Teeth cleaning prior to examination: rinsed with water Tooth drying prior to examination: not reported Threshold applied: sound 0–7; enamel 8–15; up to EDJ 16-23, beyond EDJ 24-31, dentine > 32 Device specifics: none reported
Target condition and reference standard(s)	Category: histology Sequence of index test and reference standard: index test then reference standard Training of examiner: not reported Blinding to index test: unclear Multiple tests: no Site selection: sectioned teeth Target condition: sound, enamel, dentine
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0
Comparative	
Notes	
Methodological quality	

Ouellet 2002 *(Continued)*

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		

Ouellet 2002 *(Continued)*

Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Ozsevik 2015

Study characteristics	
Patient Sampling	Method of sampling: selected Included conditions: non-cavitated and enamel lesions Teeth: permanent molars - "the teeth had no cavitations, approximal restorations, or hypoplastic pits, as judged by the naked eye" Sealants: unclear Surface: approximal
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported Country: Turkey Setting: extracted teeth, "the teeth were placed in arch models and fixed with melted utility wax. The best contact points possible were achieved" Number of participants/teeth/sites: 87 teeth/156 sites Prevalence: enamel 0.63, dentine 0.35
Index tests	Category of test: DIAGNOdent pen and Midwest, "The teeth were placed in arch models and fixed with melted utility wax. The best contact points possible were achieved" Sequence of test(s): index tests (DIAGNOdent pen and Midwest) followed by reference standard Examiner training and calibration: 1 trained examiner

Ozsevik 2015 *(Continued)*

	Teeth cleaning prior to examination: toothbrush and water (15 seconds), then 1 prophyflex Tooth drying prior to examination: air dried 3 seconds Threshold applied: sound 0–9; enamel 9.1–15; dentine > 15 Device specifics: DIAGNOdent pen: tip 1; Midwest: "red LED radiation was transported to the occlusal or approximal area using the tip of the probe in contact with the occlusal surfaces"
Target condition and reference standard(s)	Category: histology Sequence of index test and reference standard: index test then reference standard Training of examiner: calibrated Blinding to index test: yes Multiple tests: no Site selection: sectioned teeth Target condition: sound, enamel, dentine
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: 1 week to allow for separation of teeth Participants receiving both tests but excluded from results: 0
Comparative	
Notes	No evidence that the results of either index test would influence the other

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		

Ozsevik 2015 *(Continued)*

If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Ozturk 2015

Study characteristics	
Patient Sampling	Method of sampling: selected, "Selected teeth were cleaned with a rubber cup and an airwater syringe and dried for 5 sec using compressed air. Afterward, the sites were selected" Included conditions: no cavitation and early lesions Teeth: permanent molars, "Teeth with open occlusal cavities, hypoplastic fissures, occlusal restorations, occlusal fissure sealants, extensive occlusal staining, and approximal caries were excluded from the study" Sealants: unclear Surface: approximal
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported Country: Turkey Setting: extracted teeth Number of participants/teeth/sites: 44 teeth/121 sites Prevalence: enamel 0.59, dentine 0.17
Index tests	Category of test: DIAGNOdent pen Sequence of test(s): index tests (visual, radiograph - digital and cone beam, DIAGNOdent pen) then reference standard Examiner training and calibration: trained but no calibration Teeth cleaning prior to examination: rubber cup Tooth drying prior to examination: air dried, 5 seconds Threshold applied: 0-12 sounds, 13-24 initial demineralization, 25+ strong demineralization Device specifics: tip 2
Target condition and reference standard(s)	Category: histology Sequence of index test and reference standard: index test then reference standard Training of examiner: not reported Blinding to index test: not reported Multiple tests: no Site selection: sectioned teeth Target condition: sound, enamel, dentine
Flow and timing	Participants with index test but no reference standard: 0

Ozturk 2015 *(Continued)*

	Participants with reference standard but no index test: 0 Time interval between tests: 1 week to allow for separation of teeth Participants receiving both tests but excluded from results: 0
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	No		

Ozturk 2015 *(Continued)*

Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Paula 2011

Study characteristics	
Patient Sampling	Method of sampling: selected Included conditions: no cavitation and early lesions Teeth: permanent third molars Sealants: not reported Surface: occlusal
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported Country: Brazil Setting: extracted teeth - tooth bank Number of participants/teeth/sites: 26 teeth/64 teeth Prevalence: enamel 0.88, dentine 0.28

Paula 2011 *(Continued)*

Index tests	Category of test: DIAGNOdent Sequence of test(s): index tests followed by reference standard Examiner training and calibration: experienced Teeth cleaning prior to examination: pumice slurry and water Tooth drying prior to examination: air dried Threshold applied: 0-10 sound, 11-20 enamel, 21+ dentine Device specifics: tip A
Target condition and reference standard(s)	Category: histology Sequence of index test and reference standard: index test then reference standard Training of examiner: not reported Blinding to index test: unclear Multiple tests: no Site selection: sectioned teeth Target condition: sound, enamel, dentine
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: 1-2 days Participants receiving both tests but excluded from results: 0
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	

Paula 2011 *(Continued)*

Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?			
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?			
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		

Paula 2011 *(Continued)*

Did all patients receive the same reference standard?	Yes	
Were all patients included in the analysis?	Yes	
Could the patient flow have introduced bias?		Low risk

Pereira 2011

Study characteristics	
Patient Sampling	Method of sampling: selected, "None of the teeth showed macroscopic signs of cavity formation" Included conditions: no cavitation and early lesions Teeth: permanent molars Sealants: unclear Surface: occlusal
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported Country: Brazil Setting: extracted teeth Number of participants/teeth/sites: 96 teeth Prevalence: enamel 0.57, dentine 0.25
Index tests	Category of test: DIAGNOdent pen and QLF (Inspektor Research) Sequence of test(s): index tests (visual, radiograph, ECM, DIAGNOdent, QLF) then reference standard Examiner training and calibration: training event Teeth cleaning prior to examination: paste and rotating brush Tooth drying prior to examination: yes Threshold applied: DIAGNOdent: > 5 indicated caries QLF Inspektor: "The images were scored subjectively from the stored images displayed on a CRT monitor" Categories: no change in enamel fluorescence, slight change in enamel fluorescence, fluorescence loss distinctly visible without enamel broken, fluorescence loss distinctly visible with enamel broken, fluorescence loss distinctly visible with cavitation Device specifics: DIAGNOdent - tip A
Target condition and reference standard(s)	Category: histology

Pereira 2011 *(Continued)*

	Sequence of index test and reference standard: index test then reference standard Training of examiner: "Three examiners underwent a training session, which consisted of 2 h of theoretical training and 4 h of practice on extracted teeth" Blinding to index test: not reported Multiple tests: no Site selection: sectioned teeth Target condition: no caries demineralization extending to the outer ½ of the enamel demineralization extending to the inner ½ of the enamel demineralization extending to the outer ½ of the dentine demineralization extending to the outer ½ of the dentine
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: 1 week to allow for separation of teeth Participants receiving both tests but excluded from results: 0
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	Unclear		

Pereira 2011 *(Continued)*

Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Pinelli 2002

Study characteristics

Pinelli 2002 *(Continued)*

Patient Sampling	Method of sampling: selected Included conditions: no cavitation and early lesions, "none of the teeth showed macroscopic signs of cavity formation with exposure into dentin" Teeth: permanent molars, "the inclusion criterion was the presence of at least one white-spot caries lesion on a free smooth surface" Sealants: not reported Surface: "free smooth surfaces"
Patient characteristics and setting	Age: 11 to 17 years Sex: not reported Ethnicity: not reported Country: Brazil Setting: school, "The examinations were carried out in classrooms under good light conditions" Number of participants/teeth/sites: 50 participants/220 surfaces Prevalence: enamel 0.50
Index tests	Category of test: DIAGNOdent Sequence of test(s): visual then DIAGNOdent Examiner training and calibration: yes Teeth cleaning prior to examination: floss and brush Tooth drying prior to examination: air dried, 10 seconds Threshold applied: 0-4 arrested, 5+ active Device specifics: "DIAGNOdent was used to examine only the lesions detected by visual inspection"
Target condition and reference standard(s)	Category: visual Sequence of index test and reference standard: reference standard performed before index test Training of examiner: calibration completed Blinding to index test: yes Multiple tests: no Site selection: those identified visually Target condition: active or inactive lesions
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: 1 week to allow for separation of teeth

Pinelli 2002 *(Continued)*

	Participants receiving both tests but excluded from results: 0		
Comparative			
Notes	Inclusion criteria suggests sound teeth were excluded, but the results confirm that sound teeth were present in the sample		
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	No		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?			
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	No		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	

Pinelli 2002 *(Continued)*

Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	No		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Pourhashemi 2009

Study characteristics	
Patient Sampling	Method of sampling: selected Included conditions: no cavitation and early lesions, "extracted permanent premolars that seem to be intact or with primary caries in fissures" Teeth: permanent premolars Sealants: not reported Surface: occlusal
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported Country: Iran Setting: extracted teeth Number of participants/teeth/sites: 80 teeth Prevalence: not reported

Pourhashemi 2009 *(Continued)*

Index tests	Category of test: DIAGNOdent Sequence of test(s): visual, radiograph, DIAGNOdent, followed by reference standard Examiner training and calibration: not reported Teeth cleaning prior to examination: with pumice Tooth drying prior to examination: yes Threshold applied: 0-18 sound, 19-30 enamel, 30+ dentine (Lussi method) Device specifics: not reported
Target condition and reference standard(s)	Category: histology Sequence of index test and reference standard: index test then reference standard Training of examiner: not reported Blinding to index test: yes Multiple tests: no Site selection: sectioned teeth Target condition: sound, enamel, dentine
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: 1 week to allow for separation of teeth Participants receiving both tests but excluded from results: 0
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	

Pourhashemi 2009 *(Continued)*

Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		

Pourhashemi 2009 *(Continued)*

Did all patients receive the same reference standard?	Yes	
Were all patients included in the analysis?	Yes	
Could the patient flow have introduced bias?		Low risk

Presoto 2017

Study characteristics	
Patient Sampling	Method of sampling: selected Included conditions: "sound or decayed teeth" - no indication of severity of decay Teeth: permanent molars and premolars - third molars not assessed Sealants: not reported Surface: occlusal
Patient characteristics and setting	Age: young adult patients (male and female, 18 to 28 years old) Sex: not reported Ethnicity: not reported Country: Brazil Setting: clinical setting Number of participants/teeth/sites: 107 teeth/14 participants Prevalence: enamel.36
Index tests	Category of test: VistaProof Sequence of test(s): visual, radiograph, VistaProof, digital images - each assessment separated by 1 week, different examiner interpreted images Examiner training and calibration: yes - on extracted teeth Teeth cleaning prior to examination: professional prophylaxis with pumice and water Tooth drying prior to examination: drying with an air jet for 5 seconds Threshold applied: scored according to colour from heat map images, green = sound, purple = initial enamel, red = up to EDJ, orange = dentine, yellow = deep dentine Device specifics: "The results were automatically interpreted by DBSWIN software"
Target condition and reference standard(s)	Category: combined visual and radiograph Sequence of index test and reference standard: visual and radiographs performed prior to index tests but different examiners

Presoto 2017 *(Continued)*

	Training of examiner: yes Blinding to index test: yes Multiple tests: yes Site selection: visual assessment of all teeth Target condition: absence or presence of caries at enamel threshold
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: 1 week Participants receiving both tests but excluded from results: 0
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			

Presoto 2017 *(Continued)*

DOMAIN 2: Index Test (Blue fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Red fluorescence)			
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	No		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Rando-Meirelles 2011

Study characteristics	
Patient Sampling	Method of sampling: systematically selected, "the 19th child on the list was selected as the first individual of the sample, and after this every 21st child was chosen" Included conditions: not clearly reported Teeth: not clearly reported Sealants: not reported

Rando-Meirelles 2011 *(Continued)*

	Surface: not reported, assumed to be occlusal
Patient characteristics and setting	Age: 12 to 15 years Sex: not reported Ethnicity: not reported Country: Brazil Setting: school-based recruitment Number of participants/teeth/sites: 179 participants/1290 surfaces Prevalence: enamel 0.34, dentine 0.31
Index tests	Category of test: DIAGNOdent Sequence of test(s): visual, radiograph, DIAGNOdent Examiner training and calibration: trained and calibrated Teeth cleaning prior to examination: brush and paste Tooth drying prior to examination: air dried Threshold applied: 0-20 sound, 21-30 enamel, 31+ dentine Device specifics: tip A
Target condition and reference standard(s)	Category: radiograph Sequence of index test and reference standard: reference standard conducted before index test Training of examiner: trained and calibrated Blinding to index test: yes Multiple tests: no Site selection: whole tooth Target condition: sound, enamel, dentine
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: 1 week to allow for separation of teeth Participants receiving both tests but excluded from results: 0
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns

Rando-Meirelles 2011 *(Continued)*

DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	No		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	No		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	No		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		

Rando-Meirelles 2011 *(Continued)*

Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Reis 2004

Study characteristics	
Patient Sampling	Method of sampling: selected Included conditions: no cavitation and early lesions Teeth: molars and premolars Sealants: not reported Surface: occlusal
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported Country: Brazil Setting: extracted teeth Number of participants/teeth/sites: not reported Prevalence: not reported
Index tests	Category of test: DIAGNOdent Sequence of test(s): visual, radiograph, DIAGNOdent Examiner training and calibration: trained and calibrated Teeth cleaning prior to examination: yes Tooth drying prior to examination: yes Threshold applied: 0 = sound, 1 = superficial enamel, 2 = middle enamel, 3 = deep enamel, 4 = dentine

Reis 2004 *(Continued)*

	Device specifics: tip A
Target condition and reference standard(s)	Category: histology Sequence of index test and reference standard: index test conducted before reference standard Training of examiner: trained and calibrated Blinding to index test: yes Multiple tests: no Site selection: whole tooth Target condition: sound, enamel, dentine
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: 1 week to allow for separation of teeth Participants receiving both tests but excluded from results: 0
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	No		

Reis 2004 *(Continued)*

Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Reis 2006

Study characteristics	
Patient Sampling	Method of sampling: selected

Reis 2006 *(Continued)*

	Included conditions: no cavitation and early lesions Teeth: third molars Sealants: excluded Surface: occlusal
Patient characteristics and setting	Age: 19 to 30 years Sex: not reported Ethnicity: not reported Country: Brazil Setting: dental hospital Number of participants/teeth/sites: 57 teeth/110 sites Prevalence: enamel 0.82, dentine 0.15
Index tests	Category of test: DIAGNOdent Sequence of test(s): visual, DIAGNOdent Examiner training and calibration: trained Teeth cleaning prior to examination: not in clinical setting Tooth drying prior to examination: yes briefly air dried Threshold applied: 0-13 sound, 14-19 enamel or early dentine, 20+ dentinal Device specifics: tip A
Target condition and reference standard(s)	Category: histology Sequence of index test and reference standard: index test conducted before reference standard Training of examiner: not reported Blinding to index test: yes Multiple tests: no Site selection: sectioned tooth Target condition: sound, outer enamel, inner enamel and first third dentine middle and inner dentine
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: 1 week to allow for separation of teeth Participants receiving both tests but excluded from results: 0
Comparative	

Reis 2006 *(Continued)*

Notes

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			

Reis 2006 *(Continued)*

Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			High
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Ribeiro 2015

Study characteristics	
Patient Sampling	Method of sampling: selected, "The selected children presented a minimum of one tooth in an advanced stage of exfoliation" Included conditions: no cavitation and early lesions Teeth: primary molars Sealants: not reported Surface: approximal
Patient characteristics and setting	Age: 8 to 12 years Sex: not reported Ethnicity: not reported Country: Brazil Setting: school based Number of participants/teeth/sites: 112 participants/137 teeth/209 sites Prevalence: enamel 0.60, dentine 0.29
Index tests	Category of test: DIAGNOdent pen Sequence of test(s): visual, DIAGNOdent, bitewing radiograph, separators at 3 time points: baseline, 7 days later and 2 months later Examiner training and calibration: previously calibrated Teeth cleaning prior to examination: using dental floss

Ribeiro 2015 *(Continued)*

	Tooth drying prior to examination: 5 seconds with dried air Threshold applied: 0-6 sound, 6.1-9 outer enamel, 9.1-15 inner enamel, 15+ dentinal Device specifics: tip A
Target condition and reference standard(s)	Category: computed microtomography, SkyScan device (SkyScan 1174, Kontich, Belgium); description as follows: "The specimens were rotated through 360°, at a rotation speed of 1.0, with a frame average of 2 and randomized movements. A 0.25-mm aluminum filter was used. The teeth were scanned at a power of 50 kV and 800 µA"; "The teeth were three-dimensionally reconstructed using the program NRecon, version 1.6.0.3, applying maximum reduction of ring artifacts and maximum beam hardening correction (100%)" Sequence of index test and reference standard: index tests conducted before reference standard Training of examiner: not reported Blinding to index test: yes Multiple tests: no Site selection: whole tooth scanned Target condition: sound, outer enamel, inner enamel, outer dentine, inner dentine
Flow and timing	Participants with index test but no reference standard: 146 "63 surfaces out of the total sample were used for the study validation by µCT" Participants with reference standard but no index test: 0 Time interval between tests: 2 months Participants receiving both tests but excluded from results: 0
Comparative	
Notes	Primary data extracted is from the first time point prior to tooth separation as this presents the scenario closest to clinical use and is comparable to other included studies on approximal surfaces

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		

Ribeiro 2015 *(Continued)*

Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern

Ribeiro 2015 *(Continued)*

DOMAIN 4: Flow and Timing		
Was there an appropriate interval between index test and reference standard?	Yes	
Did all patients receive the same reference standard?	No	
Were all patients included in the analysis?	No	
Could the patient flow have introduced bias?		High risk

Rocha 2003

Study characteristics	
Patient Sampling	Method of sampling: selected, "Fifty occlusal sites were selected for this study" Included conditions: no cavitation and early lesions Teeth: primary molars Sealants: excluded Surface: occlusal
Patient characteristics and setting	Age: 10 to 11 years Sex: not reported Ethnicity: not reported Country: Brazil Setting: children with teeth close to exfoliation Number of participants/teeth/sites: 29 participants/50 sites Prevalence: enamel 0.58, dentine 0.14
Index tests	Category of test: DIAGNOdent Sequence of test(s): DIAGNOdent, visual, radiograph followed by reference standard Examiner training and calibration: previously trained Teeth cleaning prior to examination: "professionally cleaned" Tooth drying prior to examination: 5 seconds with dried air Threshold applied: "The cutoff limits for all and dentin lesions were values of 6 and 21" Device specifics: tip A, marked on photograph, then maximum value, mean of 3 measurements
Target condition and reference standard(s)	Category: histology

Rocha 2003 *(Continued)*

	Sequence of index test and reference standard: index test conducted before reference standard Training of examiner: not reported Blinding to index test: yes Multiple tests: no Site selection: sectioned tooth Target condition: sound, outer enamel, inner enamel and first third dentine middle and inner dentine
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern

Rocha 2003 *(Continued)*

DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Rocha-Cabral 2008

Study characteristics	
Patient Sampling	Method of sampling: selected Included conditions: no cavitation and early lesions Teeth: primary molars Sealants: not reported

Rocha-Cabral 2008 *(Continued)*

	Surface: occlusal
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported Country: Brazil Setting: extracted or recently exfoliated teeth Number of participants/teeth/sites: 66 participants/120 sites Prevalence: enamel 0.62, dentine 0.18
Index tests	Category of test: DIAGNOdent Sequence of test(s): DIAGNOdent followed by reference standard Examiner training and calibration: not reported Teeth cleaning prior to examination: "The teeth were polished with water and non-fluorescent pumice and rinsed in tap water" Tooth drying prior to examination: not reported Threshold applied: "0-4 sound/early enamel caries lesions; 5-12 advanced enamel caries; 12+ dentinal caries" Device specifics: tip A
Target condition and reference standard(s)	Category: histology Sequence of index test and reference standard: index test conducted before reference standard Training of examiner: not reported Blinding to index test: yes Multiple tests: no Site selection: sectioned tooth Target condition: sound, outer enamel, inner enamel and first third dentine middle and inner dentine
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: 1 week to allow for separation of teeth Participants receiving both tests but excluded from results: 0
Comparative	
Notes	Data not available at the relevant thresholds, includes D1 as sound, the study's primary objective was to assess the impact of autoclave on accuracy

Rocha-Cabral 2008 *(Continued)*

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			

Rocha-Cabral 2008 *(Continued)*

Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Rodrigues 2008

Study characteristics	
Patient Sampling	Method of sampling: selected Included conditions: no cavitation and early lesions Teeth: permanent molars Sealants: not reported Surface: occlusal
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported Country: Switzerland Setting: extracted teeth Number of participants/teeth/sites: 119 teeth Prevalence: enamel 0.93, dentine 0.54
Index tests	Category of test: DIAGNOdent, DIAGNOdent pen and VistaProof Sequence of test(s): DIAGNOdent, DIAGNOdent pen, VistaProof, visual, radiograph Examiner training and calibration: experienced

Rodrigues 2008 *(Continued)*

	Teeth cleaning prior to examination: yes Tooth drying prior to examination: yes Threshold applied: DIAGNOdent: 0-7 sound, 7.1-23 enamel, > 24 dentinal DIAGNOdent pen: 0-6 sound, 6.1-16 enamel, > 17 dentinal Device specifics: tip A for DIAGNOdent and cylindrical sapphire fibre tip for DIAGNOdent pen
Target condition and reference standard(s)	Category: histology Sequence of index test and reference standard: index tests conducted before reference standard Training of examiner: not reported Blinding to index test: yes Multiple tests: no Site selection: sectioned teeth Target condition: sound, enamel, outer dentine, inner dentine
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All)			

Rodrigues 2008 *(Continued)*

Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Red fluorescence)			
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		

Rodrigues 2008 *(Continued)*

Could the patient flow have introduced bias?	Low risk

Rodrigues 2009

Study characteristics	
Patient Sampling	Method of sampling: selected Included conditions: no cavitation and early lesions Teeth: primary molars Sealants: not reported Surface: occlusal
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported Country: Brazil Setting: extracted teeth Number of participants/teeth/sites: 148 teeth Prevalence: enamel 0.92, dentine 0.03
Index tests	Category of test: DIAGNOdent Sequence of test(s): visual and DIAGNOdent combined in 1 examination Examiner training and calibration: calibrated Teeth cleaning prior to examination: yes Tooth drying prior to examination: yes Threshold applied: DIAGNOdent: 0-7 sound, 7.1-23 enamel, > 24 dentinal Device specifics: tip A
Target condition and reference standard(s)	Category: histology Sequence of index test and reference standard: index tests conducted before reference standard Training of examiner: not reported Blinding to index test: yes Multiple tests: no Site selection: sectioned teeth

Rodrigues 2009 *(Continued)*

	Target condition: sound, enamel, outer dentine, inner dentine
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			

Rodrigues 2009 *(Continued)*

Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Rodrigues 2011

Study characteristics	
Patient Sampling	Method of sampling: selected Included conditions: no cavitation and early lesions Teeth: primary molars Sealants: not reported Surface: occlusal
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported

Rodrigues 2011 *(Continued)*

	Country: Switzerland Setting: extracted teeth Number of participants/teeth/sites: 97 teeth Prevalence: enamel 0.82, dentine 0.28
Index tests	Category of test: DIAGNOdent, DIAGNOdent pen, Midwest and VistaProof Sequence of test(s): index tests conducted prior to reference standard Examiner training and calibration: not reported Teeth cleaning prior to examination: yes Tooth drying prior to examination: not reported Threshold applied: DIAGNOdent and DIAGNOdent pen: not clearly reported Midwest: "change in the LED from green to red with a concurrent audible signal, confirming the presence of caries" VistaProof: calculated within study, "Optimal cut-off limits for MID and VP were determined considering the point where the sum of sensitivity and specificity was the highest" Device specifics: tip A for DIAGNOdent, cylindrical sapphire fibre tip for DIAGNOdent pen
Target condition and reference standard(s)	Category: histology Sequence of index test and reference standard: index tests conducted before reference standard Training of examiner: not reported Blinding to index test: yes Multiple tests: no Site selection: sectioned teeth Target condition: sound, enamel, outer dentine, inner dentine
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0
Comparative	
Notes	
Methodological quality	

Rodrigues 2011 *(Continued)*

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		
If multiple tests were applied were different examiners used for each (in vivo)?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		
If multiple tests were applied were different examiners used for each (in vivo)?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Red fluorescence)			
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		

Rodrigues 2011 *(Continued)*

Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Seremidi 2012

Study characteristics	
Patient Sampling	Method of sampling: selected Included conditions: no cavitation and early lesions Teeth: permanent premolars Sealants: excluded Surface: occlusal
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported Country: Greece Setting: extracted teeth Number of participants/teeth/sites: 41 teeth/107 sites Prevalence: enamel 0.78, dentine 0.19
Index tests	Category of test: DIAGNOdent pen and VistaProof Sequence of test(s): index tests (visual followed by DIAGNOdent pen and VistaProof) conducted prior to reference standard Examiner training and calibration: not reported Teeth cleaning prior to examination: yes - rubber cup and air water syringe Tooth drying prior to examination: 5 seconds compressed air

Seremidi 2012 *(Continued)*

	Threshold applied: DIAGNOdent pen: sound < 9, early enamel 9-24, deep enamel 25-44, dentine > 44 VistaProof: reported at manufacturer recommended levels and at thresholds calculated within study: manufacturer recommendations - sound < 1.3, 1.3-1.41 early enamel, 1.41-1.59 deep enamel, 1.59+ dentine; calculated within study - sound < 1, 1.0-1.49 early enamel, 1.5-1.99 deep enamel, 2.0+ dentine Device specifics: cylindrical tip for DIAGNOdent pen
Target condition and reference standard(s)	Category: histology Sequence of index test and reference standard: index tests conducted before reference standard Training of examiner: not reported Blinding to index test: not reported Multiple tests: no Site selection: sectioned teeth Target condition: sound, outer enamel, inner enamel, outer dentine, inner dentine
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0
Comparative	
Notes	Data extracted for VistaProof using manufacturer recommended thresholds, despite the thresholds calculated within study producing more accurate results. The D1+D2+D3 category was used from Table 3

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	

Seremidi 2012 *(Continued)*

Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Red fluorescence)			
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		

Seremidi 2012 *(Continued)*

Did all patients receive the same reference standard?	Yes	
Were all patients included in the analysis?	Yes	
Could the patient flow have introduced bias?		Low risk

Sheehy 2001

Study characteristics	
Patient Sampling	Method of sampling: selected Included conditions: not clearly reported, "selected for the study if a first permanent molar was erupted" Teeth: permanent first molar Sealants: not reported Surface: occlusal
Patient characteristics and setting	Age: 4.4 to 8.2 years, mean 6.85 Sex: not reported Ethnicity: not reported Country: UK Setting: unclear, but appears to be in vivo as the teeth were not extracted and sectioned Prevalence: enamel 0.55, dentine 0.28
Index tests	Category of test: DIAGNOdent Sequence of test(s): visual then DIAGNOdent Examiner training and calibration: not reported Teeth cleaning prior to examination: water and tooth-brush Tooth drying prior to examination: 3 in 1 air syringe Threshold applied: manufacturers recommendations: sound =< 14, enamel 14-20, dentine > 20 Device specifics: tapered tip
Target condition and reference standard(s)	Category: visual Sequence of index test and reference standard: reference standard performed before index test Training of examiner: not reported Blinding to index test: no Multiple tests: no

Sheehy 2001 *(Continued)*

	Site selection: "Site chosen on occlusal surface" Target condition: sound, outer enamel, inner enamel, outer dentine, inner dentine
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	No		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			

Sheehy 2001 *(Continued)*

Were the index test results interpreted without knowledge of the results of the reference standard?	No		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	No		
Were the reference standard results interpreted without knowledge of the results of the index tests?	No		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Shi 2000

Study characteristics	
Patient Sampling	Method of sampling: selected Included conditions: not clearly reported Teeth: permanent molars and premolars Sealants: no Surface: occlusal
Patient characteristics and setting	Age: not reported Sex: not reported

Shi 2000 *(Continued)*

	Ethnicity: not reported Country: Sweden Setting: extracted teeth Number of participants/teeth/sites: 76 teeth/surfaces Prevalence: enamel 0.73, dentine 0.39
Index tests	Category of test: DIAGNOdent Sequence of test(s): index then reference standard Examiner training and calibration: not reported Teeth cleaning prior to examination: yes, technique not reported Tooth drying prior to examination: yes Threshold applied: unclear - calculated within study Device specifics: conical probe
Target condition and reference standard(s)	Category: histology with microradiograph Sequence of index test and reference standard: index test before reference standard Training of examiner: not reported Blinding to index test: no Multiple tests: no Site selection: sectioned teeth according to photographed locations Target condition: sound, outer enamel, inner enamel, outer dentine, inner dentine
Flow and timing	Participants with index test but no reference standard: 6 Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		

Shi 2000 *(Continued)*

Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		
If multiple tests were applied were different examiners used for each (in vivo)?			
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		
If multiple tests were applied were different examiners used for each (in vivo)?			
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern

Shi 2000 *(Continued)*

DOMAIN 4: Flow and Timing		
Was there an appropriate interval between index test and reference standard?	Yes	
Did all patients receive the same reference standard?	Yes	
Were all patients included in the analysis?	No	
Could the patient flow have introduced bias?		High risk

Shwetha 2017

Study characteristics	
Patient Sampling	Method of sampling: unclear Included conditions: not clearly reported, "primary molars with questionable fissures that were extracted for therapeutic and orthodontic reasons" Teeth: primary molars Sealants: no Surface: occlusal
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported Country: India Setting: extracted teeth Number of participants/teeth/sites: 40 teeth/89 sites Prevalence: enamel 1.00, dentine 0.55
Index tests	Category of test: DIAGNOdent Sequence of test(s): visual, radiograph, DIAGNOdent, then reference standard Examiner training and calibration: not reported Teeth cleaning prior to examination: yes, "cleaned of all pulp remnants" Tooth drying prior to examination: not reported Threshold applied: 0-12 sound, 13-24 beginning demineralization, > 25 strong demineralization Device specifics: no tip specifics described, mean of 3 records reported
Target condition and reference standard(s)	Category: histology

Shwetha 2017 *(Continued)*

	Sequence of index test and reference standard: index test before reference standard Training of examiner: not reported Blinding to index test: not reported Multiple tests: no Site selection: sectioned teeth Target condition: enamel, or dentine
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0
Comparative	
Notes	Unable to extract data for 2x2 table as the sensitivity and specificity reported do not agree to the prevalence of disease in the text. The text states there were no sound sites (89 total sites, 43 enamel caries, 46 dentine)

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	

Shwetha 2017 *(Continued)*

Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Sinanoglu 2014

Study characteristics	
Patient Sampling	Method of sampling: unclear - referred patients, "Teeth exhibiting proximal caries in the radiological examination were excluded"

Sinanoglu 2014 *(Continued)*

	Included conditions: non-cavitated and early lesions, "Exclusion criteria for the teeth were the presence of proximal caries, surfaces that made it impossible to simulate the contact point, large carious lesions, enamel anomalies, any intrinsic or extrinsic staining, and any restorations or fissure sealants" Teeth: permanent molar and/or premolar tooth Sealants: no Surface: occlusal
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported Country: Turkey Setting: university dental school Number of participants/teeth/sites: 35 participants/217 teeth at first examination; 1 week later 11 participants/82 surfaces Prevalence: not clearly reported
Index tests	Category of test: DIAGNOdent pen Sequence of test(s): visual, radiograph, DIAGNOdent, then reference standard Examiner training and calibration: 2 experienced examiners and calibrated Teeth cleaning prior to examination: yes, "teeth were professionally cleaned using rotating brushes without any prophylactic pastes" Tooth drying prior to examination: "first examined wet and then air-dried for 5 sec" Threshold applied: 0-12 sound, 13-24 beginning demineralization, > 25 strong demineralization Device specifics: probe tip 2
Target condition and reference standard(s)	Category: excavation of severe caries, the remainder were based on a combination of visual and radiograph examinations Sequence of index test and reference standard: index test partly informs reference standard, unclear exactly how this was done Training of examiner: not reported Blinding to index test: not possible Multiple tests: yes, visual and radiograph; plus excavation Site selection: occlusal surface under investigation Target condition: no caries, enamel, or dentine
Flow and timing	Participants with index test but no reference standard: unclear how reference standard for first series of examinations in conducted, suspected that many may be missing a reference standard Participants with reference standard but no index test: 0

Sinanoglu 2014 *(Continued)*

	Time interval between tests: up to 1 week Participants receiving both tests but excluded from results: 0		
Comparative			
Notes	Unclear how Table 6 results of sensitivity and specificity are calculated, whether these are only reporting the participants that underwent excavation or a hybrid reference standard was applied to assess all participants. Also Table 9 not clear with thresholds applied and whether any sound teeth were included		
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	No		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	No		
If a threshold was used, was it pre-specified?	Yes		

Sinanoglu 2014 *(Continued)*

If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	No		
Were the reference standard results interpreted without knowledge of the results of the index tests?	No		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Unclear		
Were all patients included in the analysis?	Unclear		
Could the patient flow have introduced bias?		Unclear risk	

Souza 2013

Study characteristics	
Patient Sampling	Method of sampling: unclear Included conditions: "occlusal surfaces varying from sound to having different stages of carious lesions" Teeth: primary molars Sealants: no Surface: occlusal
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported

Souza 2013 *(Continued)*

	Country: Brazil Setting: extracted teeth Number of participants/teeth/sites: 79 teeth (42 first molars and 37 second molars) Prevalence: enamel 0.76, dentine 0.35
Index tests	Category of test: DIAGNOdent, DIAGNOdent pen, and VistaProof Sequence of test(s): visual, radiograph, DIAGNOdent, DIAGNOdent pen, and VistaProof, then reference standard; "teeth were mounted individually on a dental model" Examiner training and calibration: "Two experienced examiners independently assessed the teeth" Teeth cleaning prior to examination: yes, with sodium bicarbonate and water-powder blasting device for 10 seconds Tooth drying prior to examination: not reported Threshold applied: thresholds calculated within study: DIAGNOdent: 0-15 sound, 16-20 outer enamel, 21-30 inner enamel, > 30 dentine DIAGNOdent pen: 0-19 sound, 20-23 outer enamel, 24-35 inner enamel, > 3 dentine VistaProof: 0-1.1 sound, 1.2-1.4 outer enamel, 1.5-1.6 inner enamel, > 1.6 dentine Device specifics: tip A for the DIAGNOdent and the cylindrical sapphire fibre tip for DIAGNOdent pen. VistaProof: "The software (DBSWIN, Dürr Dental) digitised the video signal to create the occlusal surface images of 720×576 pixels with 3×8 bit intensities of RGB channels and resolution of 72 pixels/in"
Target condition and reference standard(s)	Category: histology Sequence of index test and reference standard: index test before reference standard Training of examiner: "experienced senior researcher, who did not take part in the examination" Blinding to index test: not reported Multiple tests: no Site selection: sectioned teeth Target condition: enamel or dentine
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0

Souza 2013 *(Continued)*

Comparative			
Notes			
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Unclear		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Red fluorescence)			

Souza 2013 *(Continued)*

DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Souza 2014

Study characteristics	
Patient Sampling	Method of sampling: selected Included conditions: no cavitation and early lesions Teeth: permanent upper incisors Sealants: not reported Surface: approximal
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported Country: Brazil Setting: extracted teeth Number of participants/teeth/sites: 51 teeth/102 surfaces Prevalence: enamel 0.48, dentine 0.34
Index tests	Category of test: DIAGNOdent pen, "each test tooth was placed between two sound upper incisors with a fixed position, making an anterior three-tooth group within an arch model"

Souza 2014 *(Continued)*

	Sequence of test(s): index tests (radiograph and DIAGNOdent pen (random order)) prior to reference standard Examiner training and calibration: experienced Teeth cleaning prior to examination: cleaned brush and bicarbonate Tooth drying prior to examination: unclear Threshold applied: calculated within study 0-27 sound, 28-33 enamel, 33+ dentine Device specifics: wedge shaped probe
Target condition and reference standard(s)	Category: histology Sequence of index test and reference standard: index test then reference standard Training of examiner: not reported Blinding to index test: unclear Multiple tests: no Site selection: sectioned teeth Target condition: sound, inner/outer enamel, inner/outer dentine
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High

Souza 2014 *(Continued)*

DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		

Souza 2014 *(Continued)*

Were all patients included in the analysis?	Yes
Could the patient flow have introduced bias?	Low risk

Souza 2018

Study characteristics	
Patient Sampling	Method of sampling: randomly selected Included conditions: no cavitation and early lesions (large carious lesions excluded) Teeth: primary molars Sealants: not reported Surface: approximal
Patient characteristics and setting	Age: 5 to 9 years Sex: 26 girls, 20 boys Ethnicity: not reported Country: Brazil Setting: clinical setting - dental hospital Number of participants/teeth/sites: 46 participants/195 surfaces Prevalence: enamel 0.41, dentine 0.13
Index tests	Category of test: DIAGNOdent pen Sequence of test(s): index tests (DIAGNOdent pen and radiograph) prior to reference standard Examiner training and calibration: trained and calibrated examiners Teeth cleaning prior to examination: rotating brush and floss Tooth drying prior to examination: unclear Threshold applied: 0-13 sound, 14-29 enamel, 30+ dentine (manufacturer's recommended) Device specifics: tip not reported
Target condition and reference standard(s)	Category: visual after separation Sequence of index test and reference standard: index test then reference standard Training of examiner: agreement reached between 2 examiners Blinding to index test: unclear - same examiner as index test but a week between examinations

Souza 2018 *(Continued)*

	Multiple tests: no Site selection: approximal surface after separation Target condition: sound, inner/outer enamel, inner/outer dentine		
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: 1 week Participants receiving both tests but excluded from results: 0		
Comparative			
Notes			
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			

Souza 2018 *(Continued)*

Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	No		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Sridhar 2009

Study characteristics	
Patient Sampling	Method of sampling: selected Included conditions: no cavitation and early lesions Teeth: permanent molars and premolars Sealants: no Surface: occlusal
Patient characteristics and setting	Age: not reported Sex: not reported

Sridhar 2009 *(Continued)*

	Ethnicity: not reported Country: India Setting: extracted teeth Number of participants/teeth/sites: 50 teeth Prevalence: enamel 0.96, dentine 0.12
Index tests	Category of test: DIAGNOdent Sequence of test(s): index tests (visual, radiograph, DIAGNOdent) then reference standard Examiner training and calibration: not reported Teeth cleaning prior to examination: scaled with paste Tooth drying prior to examination: air dried Threshold applied: 0-5 sound, 6-14 outer enamel, 15-20 inner enamel, 21-99 dentine Device specifics: tip A
Target condition and reference standard(s)	Category: histology Sequence of index test and reference standard: index test before reference standard Training of examiner: not reported Blinding to index test: no Multiple tests: no Site selection: sectioned teeth Target condition: sound, outer enamel, inner enamel, outer dentine, inner dentine
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 2
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		

Sridhar 2009 *(Continued)*

Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern

Sridhar 2009 *(Continued)*

DOMAIN 4: Flow and Timing		
Was there an appropriate interval between index test and reference standard?	Yes	
Did all patients receive the same reference standard?	Yes	
Were all patients included in the analysis?	Yes	
Could the patient flow have introduced bias?		Low risk

Teo 2014

Study characteristics	
Patient Sampling	Method of sampling: unclear Included conditions: no cavitation and early lesions Teeth: primary molars Sealants: no Surface: occlusal
Patient characteristics and setting	Age: 2 to 11 years Sex: not reported Ethnicity: not reported Country: UK Setting: dental school (in vivo study used, but in vitro also available) Number of participants/teeth/sites: 64 teeth/surfaces Prevalence: enamel 0.72, dentine 0.31
Index tests	Category of test: DIAGNOdent pen Sequence of test(s): index tests (visual, DIAGNOdent pen, CarieScan PRO) then reference standard Examiner training and calibration: yes on subsample Teeth cleaning prior to examination: pumice and a bristle brush Tooth drying prior to examination: not reported Threshold applied: 0-9 sound, 10-17 enamel, 18+ dentine Device specifics: not reported
Target condition and reference standard(s)	Category: histology Sequence of index test and reference standard: index test before reference standard

Teo 2014 *(Continued)*

	Training of examiner: not reported Blinding to index test: no Multiple tests: no Site selection: recorded on a drawing of the occlusal surface Target condition: sound, outer enamel, inner enamel, outer dentine, inner dentine
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern

Teo 2014 *(Continued)*

DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Tonioli 2002

Study characteristics	
Patient Sampling	Method of sampling: selected Included conditions: no cavitation and early lesions Teeth: permanent molars Sealants: no

Tonioli 2002 *(Continued)*

	Surface: occlusal
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported Country: USA Setting: extracted teeth Number of participants/teeth/sites: 29 teeth/surfaces Prevalence: enamel 0.76, dentine 0.59 (high prevalence but methods describe "early caries" as inclusion so include)
Index tests	Category of test: DIAGNOdent Sequence of test(s): index tests (visual, radiograph, DIAGNOdent) then reference standard Examiner training and calibration: yes Teeth cleaning prior to examination: scaled and polished Tooth drying prior to examination: yes but technique not reported Threshold applied: calculated from ROC Device specifics: not reported
Target condition and reference standard(s)	Category: histology Sequence of index test and reference standard: index test before reference standard Training of examiner: not reported Blinding to index test: no Multiple tests: no Site selection: sectioned teeth Target condition: sound, outer enamel, inner enamel, outer dentine, inner dentine
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 2
Comparative	
Notes	
Methodological quality	

Tonioli 2002 *(Continued)*

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		
If multiple tests were applied were different examiners used for each (in vivo)?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		
If multiple tests were applied were different examiners used for each (in vivo)?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		

Tonioli 2002 *(Continued)*

Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	No		
Could the patient flow have introduced bias?		High risk	

Tonkaboni 2018

Study characteristics	
Patient Sampling	Method of sampling: not reported Included conditions: "Teeth with large proximal cavitated carious lesions with extensive tooth destruction were excluded and replaced" Teeth: permanent molars and premolars Sealants: not reported Surface: approximal
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported Country: Iran Setting: extracted teeth Number of participants/teeth/sites: 108 teeth/324 sites reported Prevalence: contact area and higher - enamel 0.42, dentine 0.35
Index tests	Category of test: VistaCam iX, "teeth were mounted in putty impression material next to each other such that they were in contact at their marginal ridges to simulate their position in the oral cavity" Sequence of test(s): index tests (visual, radiograph, VistaCam) then reference standard Examiner training and calibration: not reported Teeth cleaning prior to examination: brushed and scaled

Tonkaboni 2018 *(Continued)*

	Tooth drying prior to examination: yes but technique not reported Threshold applied: 0 = no enamel change; IR 1 = a wide bright band with wedge-shaped structures in dark translucent enamel. The lesion may extend to the dentino-enamel junction; IR 2 = a wide bright band with wedge-shaped structures passing the dentino-enamel junction Device specifics: teeth were mounted in a putty impression. DBSWIN software was used
Target condition and reference standard(s)	Category: histology Sequence of index test and reference standard: index test before reference standard Training of examiner: not reported Blinding to index test: no Multiple tests: no Site selection: sectioned teeth Target condition: sound, outer enamel, inner enamel, outer dentine, inner dentine
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0
Comparative	
Notes	Data used from results of site at the contact area or higher

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Unclear risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		

Tonkaboni 2018 *(Continued)*

If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		Unclear risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		Unclear risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Red fluorescence)			
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Unclear risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Umemori 2010

Study characteristics	
Patient Sampling	Method of sampling: selected Included conditions: no cavitation and early lesions Teeth: permanent premolars/molars Sealants: not reported Surface: occlusal
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported Country: Japan Setting: extracted teeth Number of participants/teeth/sites: 19 participants/100 teeth Prevalence: enamel 0.36, dentine 0.12
Index tests	Category of test: DIAGNOdent Sequence of test(s): index tests (DIAGNOdent) then reference standard Examiner training and calibration: not reported Teeth cleaning prior to examination: brush and paste Tooth drying prior to examination: air dried Threshold applied: not reported Device specifics: not reported
Target condition and reference standard(s)	Category: visual - clinical diagnosis, including excavation where visual assessment warranted further investigation Sequence of index test and reference standard: index test before reference standard Training of examiner: not reported Blinding to index test: yes Multiple tests: no Site selection: unclear Target condition: sound, enamel, dentine
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: minimal

Umemori 2010 *(Continued)*

	Participants receiving both tests but excluded from results: 0		
Comparative			
Notes			
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Unclear		
If multiple tests were applied were different examiners used for each (in vivo)?			
Could the conduct or interpretation of the index test have introduced bias?		Unclear risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Unclear
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Unclear		
If multiple tests were applied were different examiners used for each (in vivo)?			
Could the conduct or interpretation of the index test have introduced bias?		Unclear risk	

Umemori 2010 *(Continued)*

Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Unclear
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	No		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	No		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		High risk	

Valera 2008

Study characteristics	
Patient Sampling	Method of sampling: selected Included conditions: no cavitation and early lesions Teeth: permanent premolars/molars Sealants: not reported Surface: occlusal
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported Country: Brazil Setting: extracted teeth Number of participants/teeth/sites: 72 teeth Prevalence: enamel 0.63, dentine 0.26
Index tests	Category of test: DIAGNOdent

Valera 2008 *(Continued)*

	Sequence of test(s): index tests (visual, radiograph, DIAGNOdent) then reference standard Examiner training and calibration: yes Teeth cleaning prior to examination: yes - sodium bicarbonate and water Tooth drying prior to examination: not reported Threshold applied: 0-5 sound, 6-10 enamel outer, 11-22 enamel deep, 21-26 dentine, 27+ deep dentine Device specifics: explorer A
Target condition and reference standard(s)	Category: histology Sequence of index test and reference standard: index test before reference standard Training of examiner: calibrated examiners Blinding to index test: yes Multiple tests: no Site selection: sectioned teeth Target condition: sound, outer enamel, inner enamel, outer dentine, inner dentine
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High

Valera 2008 *(Continued)*

DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		

Valera 2008 *(Continued)*

Were all patients included in the analysis?	Yes	
Could the patient flow have introduced bias?		Low risk

Van Hilsen 2013

Study characteristics	
Patient Sampling	Method of sampling: selected, "A single examiner sorted through collected teeth and chose an assortment of teeth without evidence of cavitated lesions (ICDAS-II 0–2)" Included conditions: no cavitation and early lesions Teeth: permanent premolars/molars Sealants: excluded Surface: occlusal
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported Country: USA Setting: extracted teeth Number of participants/teeth/sites: 45 teeth (3 damaged) Prevalence: enamel 0.76, dentine 0.31
Index tests	Category of test: Midwest Sequence of test(s): index tests (visual using digital images, fluorescence, OCT) then reference standard Examiner training and calibration: yes, "two examiners (E1, E2) were trained to use the Midwest Caries ID™ (MID) according to the manufacturer's directions" Teeth cleaning prior to examination: yes, "cleaned with pumice slurry to simulate a 'prophy cup' cleaning prior to assessment and copiously washed with water" Tooth drying prior to examination: "Teeth were kept moist" Threshold applied: sound = green/no beep, enamel = red/low frequency beep, dentine = red/high frequency beep Device specifics: "The tip of the device was inserted vertically on the surface of each tooth and moved around slightly (without pressure) in the pits and fissure area"
Target condition and reference standard(s)	Category: histology Sequence of index test and reference standard: index test before reference standard

Van Hilsen 2013 *(Continued)*

	Training of examiner: not reported Blinding to index test: yes Multiple tests: no Site selection: sectioned teeth Target condition: sound, enamel, dentine
Flow and timing	Participants with index test but no reference standard: 3 Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			

Van Hilsen 2013 *(Continued)*

DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	No		
Could the patient flow have introduced bias?		Low risk	

Virajsilp 2005

Study characteristics	
Patient Sampling	Method of sampling: selected Included conditions: no cavitation and early lesions Teeth: permanent molars Sealants: not reported Surface: approximal
Patient characteristics and setting	Age: not reported

Virajsilp 2005 *(Continued)*

	Sex: not reported Ethnicity: not reported Country: Thailand Setting: extracted teeth Number of participants/teeth/sites: 72 teeth Prevalence: enamel 0.83, dentine 0.51 (although methods state that molars without obvious cavities were recruited)
Index tests	Category of test: DIAGNOdent Sequence of test(s): index tests (visual, DIAGNOdent, radiograph) then reference standard Examiner training and calibration: yes Teeth cleaning prior to examination: scaled and polished Tooth drying prior to examination: not reported Threshold applied: calculated from ROC and not explicitly stated Device specifics: explorer A
Target condition and reference standard(s)	Category: histology Sequence of index test and reference standard: index test before reference standard Training of examiner: not reported Blinding to index test: yes Multiple tests: no Site selection: sectioned through highest DIAGNOdent value Target condition: sound, outer enamel, inner enamel, outer dentine, inner dentine
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns

Virajsilp 2005 *(Continued)*

DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	No		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	Yes		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		

Virajsilp 2005 *(Continued)*

Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Yoon 2017

Study characteristics	
Patient Sampling	Method of sampling: selected Included conditions: "large restorations, and extensive caries involving more than half of the proximal surfaces were excluded" Teeth: permanent premolars Sealants: not reported Surface: approximal
Patient characteristics and setting	Age: not reported Sex: not reported Ethnicity: not reported Country: South Korea Setting: extracted teeth Number of participants/teeth/sites: 102 teeth Prevalence: any caries 0.63, dentine level not reported
Index tests	Category of test: DIAGNOdent and QLF-D (QLF-D Biluminator 2, Inspektor Research Systems) Sequence of test(s): radiograph (reference standard) followed by QLF and DIAGNOdent Examiner training and calibration: "performed by a single skilled examiner who had sufficient training" Teeth cleaning prior to examination: "distilled water to remove soft tissue and plaque" Tooth drying prior to examination: air-dried for 5 seconds

Yoon 2017 *(Continued)*

	Threshold applied: DIAGNOdent: "value was 10 or higher" QLF: "fluorescence loss (ΔF) was measured"....."caries was diagnosed when the maximum QLF "diagnosed as caries when the fluorescence loss was lower than -13.8" Device specifics: DIAGNOdent: probe A QLF: "shutter speed of 1/15 s, aperture value of 8.0, ISO speed of 1600, white balance as manual (white light) or daylight (blue light)"
Target condition and reference standard(s)	Category: bitewing radiograph Sequence of index test and reference standard: prior to index tests Training of examiner: yes but not clearly reported Blinding to index test: done prior to index tests Multiple tests: no Site selection: not reported Target condition: sound, enamel or dentine caries
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 0
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		High risk	
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (All)			

Yoon 2017 *(Continued)*

Were the index test results interpreted without knowledge of the results of the reference standard?	No		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	No		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?	No		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Blue fluorescence)			
DOMAIN 2: Index Test (Red fluorescence)			
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	No		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		

Yoon 2017 *(Continued)*

Could the patient flow have introduced bias?	Low risk

Zeitouny 2014

Study characteristics	
Patient Sampling	Method of sampling: random Included conditions: no cavitation and early lesions Teeth: permanent molars and premolars Sealants: no Surface: occlusal
Patient characteristics and setting	Age: 15 to 65 years Sex: 11 male, 10 female Ethnicity: not reported Country: Lebanon Setting: dental school Number of participants/teeth/sites: 219 teeth Prevalence: enamel 0.74, dentine 0.14 (according to examiner 1)
Index tests	Category of test: Soprolife camera Sequence of test(s): visual (reference standard) before Soprolife Examiner training and calibration: calibrated Teeth cleaning prior to examination: waterjet and bicarbonate of soda Tooth drying prior to examination: air syringe dried for 5 seconds Threshold applied: codes 0-5: code 0 was given when the fissure appears shiny green, the enamel appears sound, and there are no visible changes; code 1 was selected if a tiny, thin red shimmer in the pits and fissure system is observed, which can slightly come up the slopes (walls) of the fissure system. No red dots appeared; code 2 darker red spots confined to the fissure are visible; code 3 dark red spots have extended as lines into the fissure areas but remain confined to the fissures. A slight beginning roughness of the more lined red areas can be visible; code 4 if the dark red (or red-orange) extends wider than the confines of the fissures; code 5 was selected if obvious openings of enamel were seen with visible dentine Device specifics: blue mode was used, Sopro imaging software was used for analysis
Target condition and reference standard(s)	Category: visual Sequence of index test and reference standard: reference standard before index test Training of examiner: trained prior to study

Zeitouny 2014 *(Continued)*

	Blinding to index test: no Multiple tests: no Site selection: unclear Target condition: ICDAS
Flow and timing	Participants with index test but no reference standard: 0 Participants with reference standard but no index test: 0 Time interval between tests: minimal Participants receiving both tests but excluded from results: 55 (see notes below)
Comparative	
Notes	Results reported for "the noncarious (sound tooth surface) lesion group that comprised the 0 scores for each method and the visual change in enamel group that included both score 1 and score 2 groups for each method." Therefore more severe levels of caries were not included in the results

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Was a case-control design avoided?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (All)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?			
Could the conduct or interpretation of the index test have introduced bias?		Low risk	

Zeitouny 2014 *(Continued)*

Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Green fluorescence)			
DOMAIN 2: Index Test (Blue fluorescence)			
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes		
If a threshold was used, was it pre-specified?	Yes		
If multiple tests were applied were different examiners used for each (in vivo)?			
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 2: Index Test (Red fluorescence)			
DOMAIN 3: Reference Standard			
Is the reference standards likely to correctly classify the target condition?	No		
Were the reference standard results interpreted without knowledge of the results of the index tests?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	No		
Could the patient flow have introduced bias?		High risk	

EDJ = enamel-dentine junction; FOTI = fibre optic transillumination; ICDAS = International Caries Detection and Assessment System; LF = laser fluorescence; OCT = optical coherence tomography; ROC = receiver operating characteristic; QLF = quantitative light-induced fluorescence.

Characteristics of excluded studies *[ordered by study ID]*

Study	Reason for exclusion
Abalos 2009	Recruited participants up to and including ICDAS 4
Abalos 2012	Recruited participants up to and including ICDAS 4
Abou 2016	Artifical caries
Abrams 2017	Thresholds used for histology do not allow for consistent classification of sound, enamel, and dentine caries used in other studies, interesting because it does use Canary system
Amaechi 2013	Used index test to inform "ground truth" so no valid reference standard
Anttonen 2004	Follow-up to the 2003 study which is included, no validation complete in this study
Askaroglou 2011	Not a DTA study, investigates correlation effects of sealants on fluorescence results
Betrisey 2014	Clear that severe caries were included in the sample
Blazejewska 2016	To be included in transillumination review as index test is DIAGNOcam
Diniz 2016	Included cavitated margins
Gomez 2013	Recruited participants up to and including ICDAS 4
Heinrich-Weltzien 2005	Study does not attempt to compare index test to a reference standard, therefore not a DTA study
Holtzman 2014	Recruited participants up to and including ICDAS 4
Jablonski-Momeni 2011a	Selected participants with "the full range of appearances from sound to gross cavitation"
Jablonski-Momeni 2013	Recruitment strategy aims to recruit dentinal lesions
Jallad 2015	Included teeth with occlusal surfaces of ICDAS 4
Kordic 2003	Table 1 confirms that dentinal caries were included in the sample
Marinova-Takorova 2014	Not a DTA study, investigates correlation only
Melo 2015	Participants were scheduled for restoration, therefore dentine decay will have been intentionally included
Menem 2017	Methods state that 30 sites were recruited with cavitated lesions, authors confirmed these to be dentinal
Mujat 2003	Not a DTA study
Mujat 2004	Not a DTA study
Nemes 2001	No suitable reference standard
Parviainen 2013	Clear from published figures that sample included frank cavitation
Patel 2014	Included lesions up to and including ICDAS 4
Pereira 2009	Same teeth and results as Pereira 2011, this paper does not report sensitivity and specificity results, instead it focusses on treatment decision

Study	Reason for exclusion
Rechmann 2012	Included lesions up to and including ICDAS 6
Subka 2019	Sample included teeth due for extraction which are described as "advanced caries"
Theocharopoulou 2015	Included frank cavitation
Zhang 2009	Root caries

DTA = diagnostic test accuracy; ICDAS = International Caries Detection and Assessment System.

D A T A

Presented below are all the data for all of the tests entered into the review.

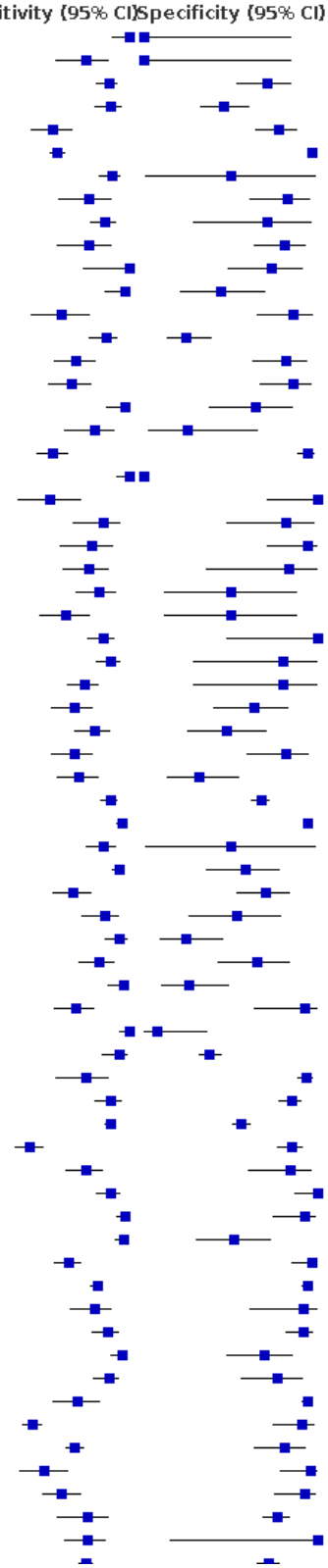
Table Tests. Data tables by test

Test	No. of studies	No. of participants
1 All	79	21283
2 Red fluorescence	68	14514
3 Blue fluorescence	20	3429
4 Green fluorescence	9	3340
5 DIAGNOdent	45	7320
6 DIAGNOdent pen	32	6842
7 VistaProof	17	2404
8 SoproLife	3	1027
9 QLF	9	3340
10 MidWest	4	356
11 Combined visual/radiograph/DIAGNOdent	1	160

Test 1. All

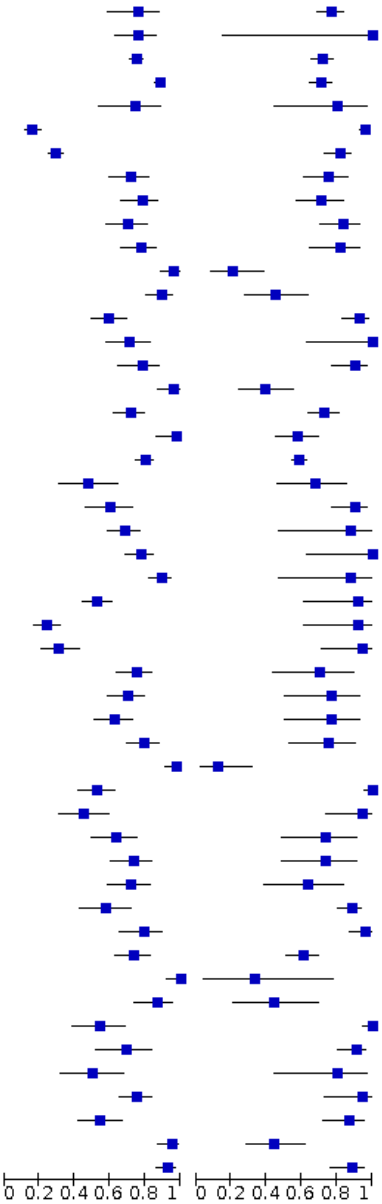
All

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Achilleos 2013	36	2	0	0	1.00 [0.90, 1.00]	0.00 [0.00, 0.84]
Achilleos 2013	27	2	9	0	0.75 [0.58, 0.88]	0.00 [0.00, 0.84]
Akarsu 2006	112	11	15	27	0.88 [0.81, 0.93]	0.71 [0.54, 0.85]
Aktan 2012	67	29	8	25	0.89 [0.80, 0.95]	0.46 [0.33, 0.60]
Aktan 2012	42	12	33	42	0.56 [0.44, 0.67]	0.78 [0.64, 0.88]
Almosa 2014	317	34	223	1079	0.59 [0.54, 0.63]	0.97 [0.96, 0.98]
Apostolopoulou 2009	98	1	11	1	0.90 [0.83, 0.95]	0.50 [0.01, 0.99]
Attrill 2001	27	4	8	19	0.77 [0.60, 0.90]	0.83 [0.61, 0.95]
Bahrololoomi 2015	88	2	14	5	0.86 [0.78, 0.92]	0.71 [0.29, 0.96]
Bamzahim 2004	26	6	8	26	0.76 [0.59, 0.89]	0.81 [0.64, 0.93]
Baseren 2003	12	5	0	14	1.00 [0.74, 1.00]	0.74 [0.49, 0.91]
Bittar 2012	36	10	1	8	0.97 [0.86, 1.00]	0.44 [0.22, 0.69]
Bittar 2012	22	3	14	19	0.61 [0.43, 0.77]	0.86 [0.65, 0.97]
Braga 2009	71	37	11	12	0.87 [0.77, 0.93]	0.24 [0.13, 0.39]
Bussaneli 2015	46	5	20	23	0.70 [0.57, 0.80]	0.82 [0.63, 0.94]
Bussaneli 2015	44	4	22	24	0.67 [0.54, 0.78]	0.86 [0.67, 0.96]
Bussaneli 2015a	41	6	1	11	0.98 [0.87, 1.00]	0.65 [0.38, 0.86]
Castilho 2016	28	6	7	2	0.80 [0.63, 0.92]	0.25 [0.03, 0.65]
Chen 2012	72	7	56	121	0.56 [0.47, 0.65]	0.95 [0.89, 0.98]
Chong 2003	48	272	0	0	1.00 [0.93, 1.00]	0.00 [0.00, 0.01]
Cinar 2013	18	0	15	11	0.55 [0.36, 0.72]	1.00 [0.72, 1.00]
Cinar 2013	28	2	5	9	0.85 [0.68, 0.95]	0.82 [0.48, 0.98]
Costa 2002	25	1	7	16	0.78 [0.60, 0.91]	0.94 [0.71, 1.00]
Diniz 2011	36	1	11	5	0.77 [0.62, 0.88]	0.83 [0.36, 1.00]
Diniz 2011	39	3	8	3	0.83 [0.69, 0.92]	0.50 [0.12, 0.88]
Diniz 2011	30	3	17	3	0.64 [0.49, 0.77]	0.50 [0.12, 0.88]
Diniz 2012	85	0	15	5	0.85 [0.76, 0.91]	1.00 [0.48, 1.00]
Diniz 2012	89	1	11	4	0.89 [0.81, 0.94]	0.80 [0.28, 0.99]
Diniz 2012	74	1	26	4	0.74 [0.64, 0.82]	0.80 [0.28, 0.99]
Diniz 2019	45	8	21	14	0.68 [0.56, 0.79]	0.64 [0.41, 0.83]
Diniz 2019	53	11	13	10	0.80 [0.69, 0.89]	0.48 [0.26, 0.70]
Diniz 2019	45	4	21	18	0.68 [0.56, 0.79]	0.82 [0.60, 0.95]
Diniz 2019	47	15	19	7	0.71 [0.59, 0.82]	0.32 [0.14, 0.55]
Duruturk 2011	163	105	20	217	0.89 [0.84, 0.93]	0.67 [0.62, 0.72]
Feng 2005	342	75	15	1300	0.96 [0.93, 0.98]	0.95 [0.93, 0.96]
Goel 2009	69	1	12	1	0.85 [0.76, 0.92]	0.50 [0.01, 0.99]
Heinrich-Weltzien 2003	212	10	12	14	0.95 [0.91, 0.97]	0.58 [0.37, 0.78]
Huth 2010	52	12	25	28	0.68 [0.56, 0.78]	0.70 [0.53, 0.83]
Iranzo-Cortes 2017	42	7	7	8	0.86 [0.73, 0.94]	0.53 [0.27, 0.79]
Jablonski-Momeni 2011	69	19	4	6	0.95 [0.87, 0.98]	0.24 [0.09, 0.45]
Jablonski-Momeni 2012	49	8	10	15	0.83 [0.71, 0.92]	0.65 [0.43, 0.84]
Jablonski-Momeni 2012	57	17	2	6	0.97 [0.88, 1.00]	0.26 [0.10, 0.48]
Jablonski-Momeni 2012a	49	1	22	12	0.69 [0.57, 0.79]	0.92 [0.64, 1.00]
Jablonski-Momeni 2012a	67	12	0	1	1.00 [0.95, 1.00]	0.08 [0.00, 0.36]
Jablonski-Momeni 2014	50	157	3	96	0.94 [0.84, 0.99]	0.38 [0.32, 0.44]
Jablonski-Momeni 2016	27	11	9	158	0.75 [0.58, 0.88]	0.93 [0.89, 0.97]
Jablonski-Momeni 2016	67	21	8	118	0.89 [0.80, 0.95]	0.85 [0.78, 0.90]
Jung 2018	333	183	39	236	0.90 [0.86, 0.92]	0.56 [0.51, 0.61]
Kim 2017	72	16	98	94	0.42 [0.35, 0.50]	0.85 [0.77, 0.91]
Ko 2015	57	3	19	16	0.75 [0.64, 0.84]	0.84 [0.60, 0.97]
Kockanat 2017	84	0	10	26	0.89 [0.81, 0.95]	1.00 [0.87, 1.00]
Kockanat 2017	92	2	2	24	0.98 [0.93, 1.00]	0.92 [0.75, 0.99]
Kouchaji 2012	129	11	4	12	0.97 [0.92, 0.99]	0.52 [0.31, 0.73]
Kucukyilmaz 2015	107	1	57	35	0.65 [0.57, 0.72]	0.97 [0.85, 1.00]
Kuhnisch 2008	489	13	110	228	0.82 [0.78, 0.85]	0.95 [0.91, 0.97]
Lee 2018	40	1	10	11	0.80 [0.66, 0.90]	0.92 [0.62, 1.00]
Lussi 2006	78	5	11	56	0.88 [0.79, 0.94]	0.92 [0.82, 0.97]
Lussi 2006a	89	8	4	18	0.96 [0.89, 0.99]	0.69 [0.48, 0.86]
Lussi 2006a	82	6	11	20	0.88 [0.80, 0.94]	0.77 [0.56, 0.91]
Mansour 2016	36	21	15	354	0.71 [0.56, 0.83]	0.94 [0.92, 0.97]
Matos 2011	155	3	195	29	0.44 [0.39, 0.50]	0.91 [0.75, 0.98]
Matos 2011	240	6	110	26	0.69 [0.63, 0.73]	0.81 [0.64, 0.93]
Mendes 2005	27	1	26	23	0.51 [0.37, 0.65]	0.96 [0.79, 1.00]
Mendes 2006	51	2	32	25	0.61 [0.50, 0.72]	0.93 [0.76, 0.99]
Mepparambath 2014	28	31	9	101	0.76 [0.59, 0.88]	0.77 [0.68, 0.83]
Mortensen 2018	44	0	14	2	0.76 [0.63, 0.86]	1.00 [0.16, 1.00]
Muller-Bolla 2017	404	58	122	148	0.75 [0.71, 0.79]	0.72 [0.65, 0.78]



Test 1. (Continued)

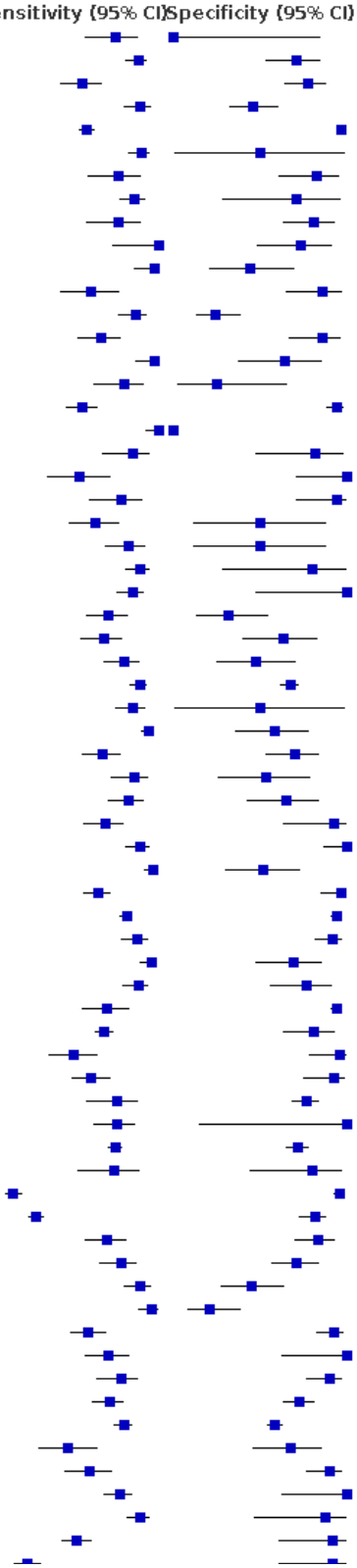
Mepparambath 2014	28	31	9	101	0.76 [0.59, 0.88]	0.77 [0.68, 0.83]
Mortensen 2018	44	0	14	2	0.76 [0.63, 0.86]	1.00 [0.16, 1.00]
Muller-Bolla 2017	404	58	133	148	0.75 [0.71, 0.79]	0.72 [0.65, 0.78]
Muller-Bolla 2017	473	60	62	148	0.88 [0.85, 0.91]	0.71 [0.64, 0.77]
Neuhaus 2011	20	2	7	8	0.74 [0.54, 0.89]	0.80 [0.44, 0.97]
Novaes 2009	41	15	215	350	0.16 [0.12, 0.21]	0.96 [0.93, 0.98]
Novaes 2010	141	21	337	93	0.29 [0.25, 0.34]	0.82 [0.73, 0.88]
Novaes 2012a	46	12	18	37	0.72 [0.59, 0.82]	0.76 [0.61, 0.87]
Novaes 2012a	50	14	14	35	0.78 [0.66, 0.87]	0.71 [0.57, 0.83]
Novaes 2012a	45	8	19	41	0.70 [0.58, 0.81]	0.84 [0.70, 0.93]
Novaes 2016	59	6	17	27	0.78 [0.67, 0.86]	0.82 [0.65, 0.93]
Novaes 2016	73	26	3	7	0.96 [0.89, 0.99]	0.21 [0.09, 0.39]
Novaes 2016	68	18	8	15	0.89 [0.80, 0.95]	0.45 [0.28, 0.64]
Ozsevik 2015	59	4	40	53	0.60 [0.49, 0.69]	0.93 [0.83, 0.98]
Paula 2011	40	0	16	8	0.71 [0.58, 0.83]	1.00 [0.63, 1.00]
Pereira 2011	43	4	12	37	0.78 [0.65, 0.88]	0.90 [0.77, 0.97]
Pereira 2011	53	25	2	16	0.96 [0.87, 1.00]	0.39 [0.24, 0.55]
Pinelli 2002	78	30	31	81	0.72 [0.62, 0.80]	0.73 [0.64, 0.81]
Presoto 2017	37	29	1	40	0.97 [0.86, 1.00]	0.58 [0.45, 0.70]
Rando-Meirelles 2011	213	215	53	308	0.80 [0.75, 0.85]	0.59 [0.55, 0.63]
Ribeiro 2015	18	8	20	17	0.47 [0.31, 0.64]	0.68 [0.46, 0.85]
Rocha 2003	35	4	23	38	0.60 [0.47, 0.73]	0.90 [0.77, 0.97]
Rodrigues 2008	76	1	35	7	0.68 [0.59, 0.77]	0.88 [0.47, 1.00]
Rodrigues 2008	86	0	25	8	0.77 [0.69, 0.85]	1.00 [0.63, 1.00]
Rodrigues 2008	99	1	12	7	0.89 [0.82, 0.94]	0.88 [0.47, 1.00]
Rodrigues 2009	83	1	74	11	0.53 [0.45, 0.61]	0.92 [0.62, 1.00]
Rodrigues 2009	33	1	103	11	0.24 [0.17, 0.32]	0.92 [0.62, 1.00]
Rodrigues 2011	25	1	55	16	0.31 [0.21, 0.43]	0.94 [0.71, 1.00]
Rodrigues 2011	60	5	20	12	0.75 [0.64, 0.84]	0.71 [0.44, 0.90]
Rodrigues 2011	56	4	24	13	0.70 [0.59, 0.80]	0.76 [0.50, 0.93]
Rodrigues 2011	50	4	30	13	0.63 [0.51, 0.73]	0.76 [0.50, 0.93]
Seremidi 2012	66	6	17	18	0.80 [0.69, 0.88]	0.75 [0.53, 0.90]
Seremidi 2012	81	21	2	3	0.98 [0.92, 1.00]	0.13 [0.03, 0.32]
Sheehy 2001	49	0	44	77	0.53 [0.42, 0.63]	1.00 [0.95, 1.00]
Shi 2000	23	1	28	18	0.45 [0.31, 0.60]	0.95 [0.74, 1.00]
Souza 2013	38	5	22	14	0.63 [0.50, 0.75]	0.74 [0.49, 0.91]
Souza 2013	44	5	16	14	0.73 [0.60, 0.84]	0.74 [0.49, 0.91]
Souza 2013	43	7	17	12	0.72 [0.59, 0.83]	0.63 [0.38, 0.84]
Souza 2014	29	11	21	83	0.58 [0.43, 0.72]	0.88 [0.80, 0.94]
Souza 2014	39	2	10	51	0.80 [0.66, 0.90]	0.96 [0.87, 1.00]
Souza 2018	59	45	21	70	0.74 [0.63, 0.83]	0.61 [0.51, 0.70]
Sridhar 2009	44	4	0	2	1.00 [0.92, 1.00]	0.33 [0.04, 0.78]
Teo 2014	40	10	6	8	0.87 [0.74, 0.95]	0.44 [0.22, 0.69]
Tonkaboni 2018	25	0	21	62	0.54 [0.39, 0.69]	1.00 [0.94, 1.00]
Umemori 2010	25	6	11	58	0.69 [0.52, 0.84]	0.91 [0.81, 0.96]
Van Hilsen 2013	16	2	16	8	0.50 [0.32, 0.68]	0.80 [0.44, 0.97]
Virajsilp 2005	67	1	22	17	0.75 [0.65, 0.84]	0.94 [0.73, 1.00]
Yoon 2017	35	5	29	33	0.55 [0.42, 0.67]	0.87 [0.72, 0.96]
Yoon 2017	61	21	3	17	0.95 [0.87, 0.99]	0.45 [0.29, 0.62]
Zeitouny 2014	104	6	8	46	0.93 [0.86, 0.97]	0.88 [0.77, 0.96]



Test 2. Red fluorescence

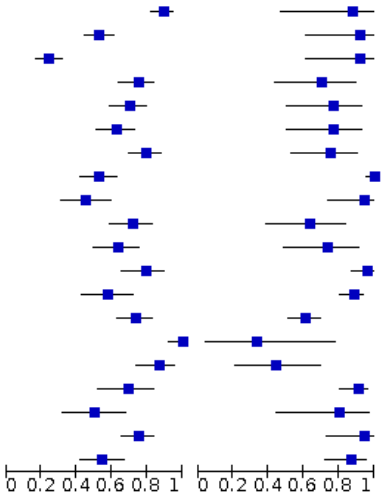
Red fluorescence

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Achilleos 2013	27	2	9	0	0.75 [0.58, 0.88]	0.00 [0.00, 0.84]
Akarsu 2006	112	11	15	27	0.88 [0.81, 0.93]	0.71 [0.54, 0.85]
Aktan 2012	42	12	33	42	0.56 [0.44, 0.67]	0.78 [0.64, 0.88]
Aktan 2012	67	29	8	25	0.89 [0.80, 0.95]	0.46 [0.33, 0.60]
Almosa 2014	317	34	223	1079	0.59 [0.54, 0.63]	0.97 [0.96, 0.98]
Apostolopoulou 2009	98	1	11	1	0.90 [0.83, 0.95]	0.50 [0.01, 0.99]
Attrill 2001	27	4	8	19	0.77 [0.60, 0.90]	0.83 [0.61, 0.95]
Bahrololoomi 2015	88	2	14	5	0.86 [0.78, 0.92]	0.71 [0.29, 0.96]
Bamzahim 2004	26	6	8	26	0.76 [0.59, 0.89]	0.81 [0.64, 0.93]
Baseren 2003	12	5	0	14	1.00 [0.74, 1.00]	0.74 [0.49, 0.91]
Bittar 2012	36	10	1	8	0.97 [0.86, 1.00]	0.44 [0.22, 0.69]
Bittar 2012	22	3	14	19	0.61 [0.43, 0.77]	0.86 [0.65, 0.97]
Braga 2009	71	37	11	12	0.87 [0.77, 0.93]	0.24 [0.13, 0.39]
Bussaneli 2015	44	4	22	24	0.67 [0.54, 0.78]	0.86 [0.67, 0.96]
Bussaneli 2015a	41	6	1	11	0.98 [0.87, 1.00]	0.65 [0.38, 0.86]
Castilho 2016	28	6	7	2	0.80 [0.63, 0.92]	0.25 [0.03, 0.65]
Chen 2012	72	7	56	121	0.56 [0.47, 0.65]	0.95 [0.89, 0.98]
Chong 2003	48	272	0	0	1.00 [0.93, 1.00]	0.00 [0.00, 0.01]
Cinar 2013	28	2	5	9	0.85 [0.68, 0.95]	0.82 [0.48, 0.98]
Cinar 2013	18	0	15	11	0.55 [0.36, 0.72]	1.00 [0.72, 1.00]
Costa 2002	25	1	7	16	0.78 [0.60, 0.91]	0.94 [0.71, 1.00]
Diniz 2011	30	3	17	3	0.64 [0.49, 0.77]	0.50 [0.12, 0.88]
Diniz 2011	39	3	8	3	0.83 [0.69, 0.92]	0.50 [0.12, 0.88]
Diniz 2012	89	1	11	4	0.89 [0.81, 0.94]	0.80 [0.28, 0.99]
Diniz 2012	85	0	15	5	0.85 [0.76, 0.91]	1.00 [0.48, 1.00]
Diniz 2019	47	15	19	7	0.71 [0.59, 0.82]	0.32 [0.14, 0.55]
Diniz 2019	45	8	21	14	0.68 [0.56, 0.79]	0.64 [0.41, 0.83]
Diniz 2019	53	11	13	10	0.80 [0.69, 0.89]	0.48 [0.26, 0.70]
Duruturk 2011	163	105	20	217	0.89 [0.84, 0.93]	0.67 [0.62, 0.72]
Goel 2009	69	1	12	1	0.85 [0.76, 0.92]	0.50 [0.01, 0.99]
Heinrich-Weltzien 2003	212	10	12	14	0.95 [0.91, 0.97]	0.58 [0.37, 0.78]
Huth 2010	52	12	25	28	0.68 [0.56, 0.78]	0.70 [0.53, 0.83]
Iranzo-Cortes 2017	42	7	7	8	0.86 [0.73, 0.94]	0.53 [0.27, 0.79]
Jablonski-Momeni 2012	49	8	10	15	0.83 [0.71, 0.92]	0.65 [0.43, 0.84]
Jablonski-Momeni 2012a	49	1	22	12	0.69 [0.57, 0.79]	0.92 [0.64, 1.00]
Kockanat 2017	84	0	10	26	0.89 [0.81, 0.95]	1.00 [0.87, 1.00]
Kouchaji 2012	129	11	4	12	0.97 [0.92, 0.99]	0.52 [0.31, 0.73]
Kucukyilmaz 2015	107	1	57	35	0.65 [0.57, 0.72]	0.97 [0.85, 1.00]
Kuhnisch 2008	489	13	110	228	0.82 [0.78, 0.85]	0.95 [0.91, 0.97]
Lussi 2006	78	5	11	56	0.88 [0.79, 0.94]	0.92 [0.82, 0.97]
Lussi 2006a	89	8	4	18	0.96 [0.89, 0.99]	0.69 [0.48, 0.86]
Lussi 2006a	82	6	11	20	0.88 [0.80, 0.94]	0.77 [0.56, 0.91]
Mansour 2016	36	21	15	354	0.71 [0.56, 0.83]	0.94 [0.92, 0.97]
Matos 2011	240	6	110	26	0.69 [0.63, 0.73]	0.81 [0.64, 0.93]
Mendes 2005	27	1	26	23	0.51 [0.37, 0.65]	0.96 [0.79, 1.00]
Mendes 2006	51	2	32	25	0.61 [0.50, 0.72]	0.93 [0.76, 0.99]
Mepparambath 2014	28	31	9	101	0.76 [0.59, 0.88]	0.77 [0.68, 0.83]
Mortensen 2018	44	0	14	2	0.76 [0.63, 0.86]	1.00 [0.16, 1.00]
Muller-Bolla 2017	404	58	133	148	0.75 [0.71, 0.79]	0.72 [0.65, 0.78]
Neuhaus 2011	20	2	7	8	0.74 [0.54, 0.89]	0.80 [0.44, 0.97]
Novaes 2009	41	15	215	350	0.16 [0.12, 0.21]	0.96 [0.93, 0.98]
Novaes 2010	141	21	337	93	0.29 [0.25, 0.34]	0.82 [0.73, 0.88]
Novaes 2012a	45	8	19	41	0.70 [0.58, 0.81]	0.84 [0.70, 0.93]
Novaes 2012a	50	14	14	35	0.78 [0.66, 0.87]	0.71 [0.57, 0.83]
Novaes 2016	68	18	8	15	0.89 [0.80, 0.95]	0.45 [0.28, 0.64]
Novaes 2016	73	26	3	7	0.96 [0.89, 0.99]	0.21 [0.09, 0.39]
Ozsevik 2015	59	4	40	53	0.60 [0.49, 0.69]	0.93 [0.83, 0.98]
Paula 2011	40	0	16	8	0.71 [0.58, 0.83]	1.00 [0.63, 1.00]
Pereira 2011	43	4	12	37	0.78 [0.65, 0.88]	0.90 [0.77, 0.97]
Pinelli 2002	78	30	31	81	0.72 [0.62, 0.80]	0.73 [0.64, 0.81]
Rando-Meirelles 2011	213	215	53	308	0.80 [0.75, 0.85]	0.59 [0.55, 0.63]
Ribeiro 2015	18	8	20	17	0.47 [0.31, 0.64]	0.68 [0.46, 0.85]
Rocha 2003	35	4	23	38	0.60 [0.47, 0.73]	0.90 [0.77, 0.97]
Rodrigues 2008	86	0	25	8	0.77 [0.69, 0.85]	1.00 [0.63, 1.00]
Rodrigues 2008	99	1	12	7	0.89 [0.82, 0.94]	0.88 [0.47, 1.00]
Rodrigues 2009	83	1	74	11	0.53 [0.45, 0.61]	0.92 [0.62, 1.00]



Test 2. (Continued)

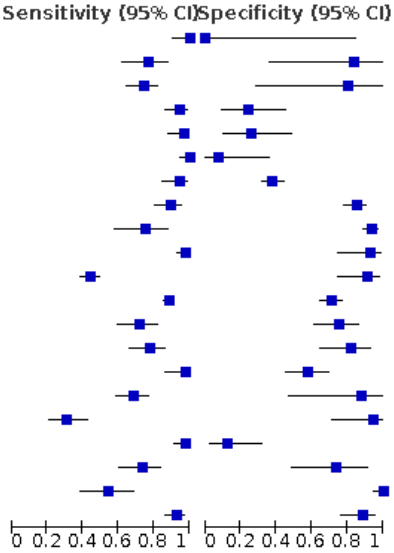
Rodrigues 2008	99	1	12	7	0.89 [0.82, 0.94]	0.88 [0.47, 1.00]
Rodrigues 2009	83	1	74	11	0.53 [0.45, 0.61]	0.92 [0.62, 1.00]
Rodrigues 2009	33	1	103	11	0.24 [0.17, 0.32]	0.92 [0.62, 1.00]
Rodrigues 2011	60	5	20	12	0.75 [0.64, 0.84]	0.71 [0.44, 0.90]
Rodrigues 2011	56	4	24	13	0.70 [0.59, 0.80]	0.76 [0.50, 0.93]
Rodrigues 2011	50	4	30	13	0.63 [0.51, 0.73]	0.76 [0.50, 0.93]
Seremidi 2012	66	6	17	18	0.80 [0.69, 0.88]	0.75 [0.53, 0.90]
Sheehy 2001	49	0	44	77	0.53 [0.42, 0.63]	1.00 [0.95, 1.00]
Shi 2000	23	1	28	18	0.45 [0.31, 0.60]	0.95 [0.74, 1.00]
Souza 2013	43	7	17	12	0.72 [0.59, 0.83]	0.63 [0.38, 0.84]
Souza 2013	38	5	22	14	0.63 [0.50, 0.75]	0.74 [0.49, 0.91]
Souza 2014	39	2	10	51	0.80 [0.66, 0.90]	0.96 [0.87, 1.00]
Souza 2014	29	11	21	83	0.58 [0.43, 0.72]	0.88 [0.80, 0.94]
Souza 2018	59	45	21	70	0.74 [0.63, 0.83]	0.61 [0.51, 0.70]
Sridhar 2009	44	4	0	2	1.00 [0.92, 1.00]	0.33 [0.04, 0.78]
Teo 2014	40	10	6	8	0.87 [0.74, 0.95]	0.44 [0.22, 0.69]
Umemori 2010	25	6	11	58	0.69 [0.52, 0.84]	0.91 [0.81, 0.96]
Van Hilsen 2013	16	2	16	8	0.50 [0.32, 0.68]	0.80 [0.44, 0.97]
Virajsilp 2005	67	1	22	17	0.75 [0.65, 0.84]	0.94 [0.73, 1.00]
Yoon 2017	35	5	29	33	0.55 [0.42, 0.67]	0.87 [0.72, 0.96]



Test 3. Blue fluorescence

Blue fluorescence

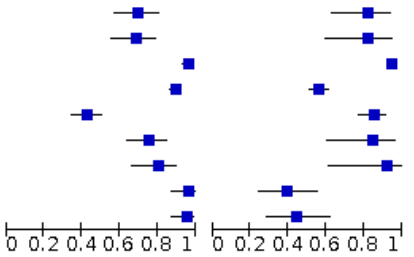
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Achilleos 2013	36	2	0	0	1.00 [0.90, 1.00]	0.00 [0.00, 0.84]
Diniz 2011	36	1	11	5	0.77 [0.62, 0.88]	0.83 [0.36, 1.00]
Diniz 2012	74	1	26	4	0.74 [0.64, 0.82]	0.80 [0.28, 0.99]
Jablonski-Momeni 2011	69	19	4	6	0.95 [0.87, 0.98]	0.24 [0.09, 0.45]
Jablonski-Momeni 2012	57	17	2	6	0.97 [0.88, 1.00]	0.26 [0.10, 0.48]
Jablonski-Momeni 2012a	67	12	0	1	1.00 [0.95, 1.00]	0.08 [0.00, 0.36]
Jablonski-Momeni 2014	50	157	3	96	0.94 [0.84, 0.99]	0.38 [0.32, 0.44]
Jablonski-Momeni 2016	67	21	8	118	0.89 [0.80, 0.95]	0.85 [0.78, 0.90]
Jablonski-Momeni 2016	27	11	9	158	0.75 [0.58, 0.88]	0.93 [0.89, 0.97]
Kockanat 2017	92	2	2	24	0.98 [0.93, 1.00]	0.92 [0.75, 0.99]
Matos 2011	155	3	195	29	0.44 [0.39, 0.50]	0.91 [0.75, 0.98]
Muller-Bolla 2017	473	60	62	148	0.88 [0.85, 0.91]	0.71 [0.64, 0.77]
Novaes 2012a	46	12	18	37	0.72 [0.59, 0.82]	0.76 [0.61, 0.87]
Novaes 2016	59	6	17	27	0.78 [0.67, 0.86]	0.82 [0.65, 0.93]
Presoto 2017	37	29	1	40	0.97 [0.86, 1.00]	0.58 [0.45, 0.70]
Rodrigues 2008	76	1	35	7	0.68 [0.59, 0.77]	0.88 [0.47, 1.00]
Rodrigues 2011	25	1	55	16	0.31 [0.21, 0.43]	0.94 [0.71, 1.00]
Seremidi 2012	81	21	2	3	0.98 [0.92, 1.00]	0.13 [0.03, 0.32]
Souza 2013	44	5	16	14	0.73 [0.60, 0.84]	0.74 [0.49, 0.91]
Tonkaboni 2018	25	0	21	62	0.54 [0.39, 0.69]	1.00 [0.94, 1.00]
Zeitouny 2014	104	6	8	46	0.93 [0.86, 0.97]	0.88 [0.77, 0.96]



Test 4. Green fluorescence

Green fluorescence

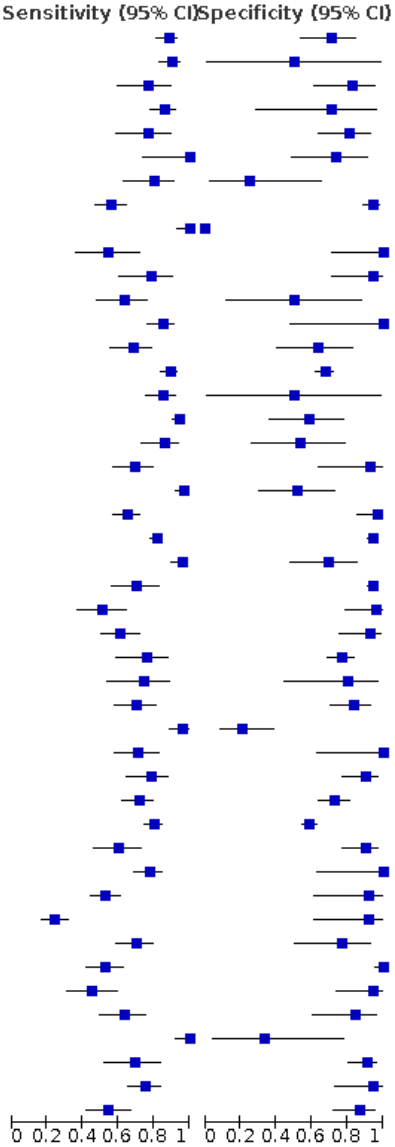
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Bussaneli 2015	46	5	20	23	0.70 [0.57, 0.80]	0.82 [0.63, 0.94]
Diniz 2019	45	4	21	18	0.68 [0.56, 0.79]	0.82 [0.60, 0.95]
Feng 2005	342	75	15	1300	0.96 [0.93, 0.98]	0.95 [0.93, 0.96]
Jung 2018	333	183	39	236	0.90 [0.86, 0.92]	0.56 [0.51, 0.61]
Kim 2017	72	16	98	94	0.42 [0.35, 0.50]	0.85 [0.77, 0.91]
Ko 2015	57	3	19	16	0.75 [0.64, 0.84]	0.84 [0.60, 0.97]
Lee 2018	40	1	10	11	0.80 [0.66, 0.90]	0.92 [0.62, 1.00]
Pereira 2011	53	25	2	16	0.96 [0.87, 1.00]	0.39 [0.24, 0.55]
Yoon 2017	61	21	3	17	0.95 [0.87, 0.99]	0.45 [0.29, 0.62]



Test 5. DIAGNOdent

DIAGNOdent

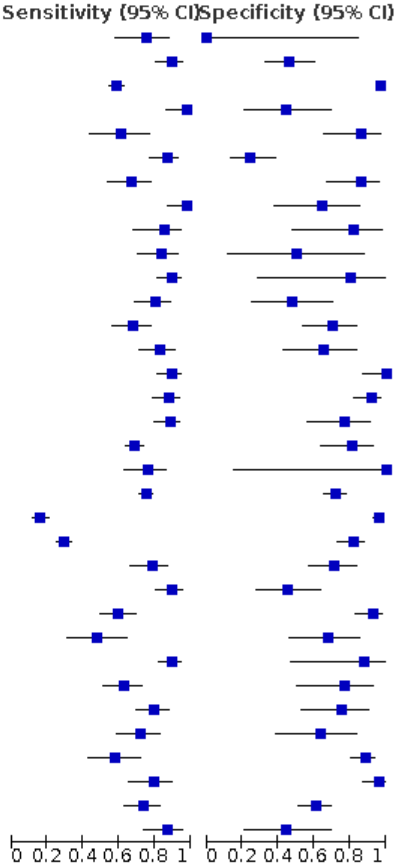
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Akarsu 2006	112	11	15	27	0.88 [0.81, 0.93]	0.71 [0.54, 0.85]
Apostolopoulou 2009	98	1	11	1	0.90 [0.83, 0.95]	0.50 [0.01, 0.99]
Attrill 2001	27	4	8	19	0.77 [0.60, 0.90]	0.83 [0.61, 0.95]
Bahrololoomi 2015	88	2	14	5	0.86 [0.78, 0.92]	0.71 [0.29, 0.96]
Bamzahim 2004	26	6	8	26	0.76 [0.59, 0.89]	0.81 [0.64, 0.93]
Baseren 2003	12	5	0	14	1.00 [0.74, 1.00]	0.74 [0.49, 0.91]
Castilho 2016	28	6	7	2	0.80 [0.63, 0.92]	0.25 [0.03, 0.65]
Chen 2012	72	7	56	121	0.56 [0.47, 0.65]	0.95 [0.89, 0.98]
Chong 2003	48	272	0	0	1.00 [0.93, 1.00]	0.00 [0.00, 0.01]
Cinar 2013	18	0	15	11	0.55 [0.36, 0.72]	1.00 [0.72, 1.00]
Costa 2002	25	1	7	16	0.78 [0.60, 0.91]	0.94 [0.71, 1.00]
Diniz 2011	31	3	18	3	0.63 [0.48, 0.77]	0.50 [0.12, 0.88]
Diniz 2012	85	0	15	5	0.85 [0.76, 0.91]	1.00 [0.48, 1.00]
Diniz 2019	45	8	21	14	0.68 [0.56, 0.79]	0.64 [0.41, 0.83]
Duruturk 2011	163	105	20	217	0.89 [0.84, 0.93]	0.67 [0.62, 0.72]
Goel 2009	69	1	12	1	0.85 [0.76, 0.92]	0.50 [0.01, 0.99]
Heinrich-Weltzien 2003	212	10	12	14	0.95 [0.91, 0.97]	0.58 [0.37, 0.78]
Iranzo-Cortes 2017	42	7	7	8	0.86 [0.73, 0.94]	0.53 [0.27, 0.79]
Jablonski-Momeni 2012a	49	1	22	12	0.69 [0.57, 0.79]	0.92 [0.64, 1.00]
Kouchaji 2012	129	11	4	12	0.97 [0.92, 0.99]	0.52 [0.31, 0.73]
Kucukyilmaz 2015	107	1	57	35	0.65 [0.57, 0.72]	0.97 [0.85, 1.00]
Kuhnisch 2008	489	13	110	228	0.82 [0.78, 0.85]	0.95 [0.91, 0.97]
Lussi 2006a	89	8	4	18	0.96 [0.89, 0.99]	0.69 [0.48, 0.86]
Mansour 2016	36	21	15	354	0.71 [0.56, 0.83]	0.94 [0.92, 0.97]
Mendes 2005	27	1	26	23	0.51 [0.37, 0.65]	0.96 [0.79, 1.00]
Mendes 2006	51	2	32	25	0.61 [0.50, 0.72]	0.93 [0.76, 0.99]
Mepparambath 2014	28	31	9	101	0.76 [0.59, 0.88]	0.77 [0.68, 0.83]
Neuhaus 2011	20	2	7	8	0.74 [0.54, 0.89]	0.80 [0.44, 0.97]
Novaes 2012a	45	8	19	41	0.70 [0.58, 0.81]	0.84 [0.70, 0.93]
Novaes 2016	73	26	3	7	0.96 [0.89, 0.99]	0.21 [0.09, 0.39]
Paula 2011	40	0	16	8	0.71 [0.58, 0.83]	1.00 [0.63, 1.00]
Pereira 2011	43	4	12	37	0.78 [0.65, 0.88]	0.90 [0.77, 0.97]
Pinelli 2002	78	30	31	81	0.72 [0.62, 0.80]	0.73 [0.64, 0.81]
Rando-Meirelles 2011	213	215	53	308	0.80 [0.75, 0.85]	0.59 [0.55, 0.63]
Rocha 2003	35	4	23	38	0.60 [0.47, 0.73]	0.90 [0.77, 0.97]
Rodrigues 2008	86	0	25	8	0.77 [0.69, 0.85]	1.00 [0.63, 1.00]
Rodrigues 2009	83	1	74	11	0.53 [0.45, 0.61]	0.92 [0.62, 1.00]
Rodrigues 2009	33	1	103	11	0.24 [0.17, 0.32]	0.92 [0.62, 1.00]
Rodrigues 2011	56	4	24	13	0.70 [0.59, 0.80]	0.76 [0.50, 0.93]
Sheehy 2001	49	0	44	77	0.53 [0.42, 0.63]	1.00 [0.95, 1.00]
Shi 2000	23	1	28	18	0.45 [0.31, 0.60]	0.95 [0.74, 1.00]
Souza 2013	38	3	22	16	0.63 [0.50, 0.75]	0.84 [0.60, 0.97]
Sridhar 2009	44	4	0	2	1.00 [0.92, 1.00]	0.33 [0.04, 0.78]
Umemori 2010	25	6	11	58	0.69 [0.52, 0.84]	0.91 [0.81, 0.96]
Virajsilp 2005	67	1	22	17	0.75 [0.65, 0.84]	0.94 [0.73, 1.00]
Yoon 2017	35	5	29	33	0.55 [0.42, 0.67]	0.87 [0.72, 0.96]



Test 6. DIAGNOdent pen

DIAGNOdent pen

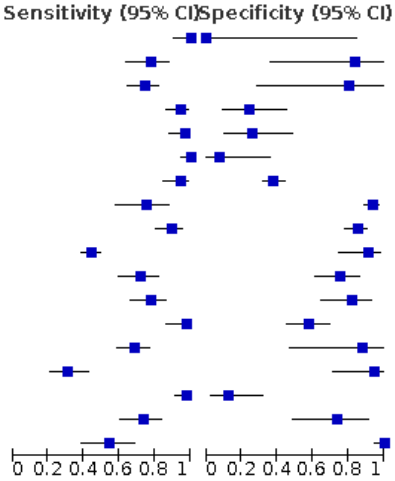
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Achilleos 2013	27	2	9	0	0.75 [0.58, 0.88]	0.00 [0.00, 0.84]
Aktan 2012	67	29	8	25	0.89 [0.80, 0.95]	0.46 [0.33, 0.60]
Almosa 2014	317	34	223	1079	0.59 [0.54, 0.63]	0.97 [0.96, 0.98]
Bittar 2012	36	10	1	8	0.97 [0.86, 1.00]	0.44 [0.22, 0.69]
Bittar 2012	22	3	14	19	0.61 [0.43, 0.77]	0.86 [0.65, 0.97]
Braga 2009	71	37	11	12	0.87 [0.77, 0.93]	0.24 [0.13, 0.39]
Bussaneli 2015	44	4	22	24	0.67 [0.54, 0.78]	0.86 [0.67, 0.96]
Bussaneli 2015a	41	6	1	11	0.98 [0.87, 1.00]	0.65 [0.38, 0.86]
Cinar 2013	28	2	5	9	0.85 [0.68, 0.95]	0.82 [0.48, 0.98]
Diniz 2011	41	3	8	3	0.84 [0.70, 0.93]	0.50 [0.12, 0.88]
Diniz 2012	89	1	11	4	0.89 [0.81, 0.94]	0.80 [0.28, 0.99]
Diniz 2019	53	11	13	10	0.80 [0.69, 0.89]	0.48 [0.26, 0.70]
Huth 2010	52	12	25	28	0.68 [0.56, 0.78]	0.70 [0.53, 0.83]
Jablonski-Momeni 2012	49	8	10	15	0.83 [0.71, 0.92]	0.65 [0.43, 0.84]
Kockanat 2017	84	0	10	26	0.89 [0.81, 0.95]	1.00 [0.87, 1.00]
Lussi 2006	78	5	11	56	0.88 [0.79, 0.94]	0.92 [0.82, 0.97]
Lussi 2006a	82	6	11	20	0.88 [0.80, 0.94]	0.77 [0.56, 0.91]
Matos 2011	240	6	110	26	0.69 [0.63, 0.73]	0.81 [0.64, 0.93]
Mortensen 2018	44	0	14	2	0.76 [0.63, 0.86]	1.00 [0.16, 1.00]
Muller-Bolla 2017	404	58	133	148	0.75 [0.71, 0.79]	0.72 [0.65, 0.78]
Novaes 2009	41	15	215	350	0.16 [0.12, 0.21]	0.96 [0.93, 0.98]
Novaes 2010	141	21	337	93	0.29 [0.25, 0.34]	0.82 [0.73, 0.88]
Novaes 2012a	50	14	14	35	0.78 [0.66, 0.87]	0.71 [0.57, 0.83]
Novaes 2016	68	18	8	15	0.89 [0.80, 0.95]	0.45 [0.28, 0.64]
Ozsevik 2015	59	4	40	53	0.60 [0.49, 0.69]	0.93 [0.83, 0.98]
Ribeiro 2015	18	8	20	17	0.47 [0.31, 0.64]	0.68 [0.46, 0.85]
Rodrigues 2008	99	1	12	7	0.89 [0.82, 0.94]	0.88 [0.47, 1.00]
Rodrigues 2011	50	4	30	13	0.63 [0.51, 0.73]	0.76 [0.50, 0.93]
Seremidi 2012	66	6	17	18	0.80 [0.69, 0.88]	0.75 [0.53, 0.90]
Souza 2013	43	7	17	12	0.72 [0.59, 0.83]	0.63 [0.38, 0.84]
Souza 2014	29	11	21	83	0.58 [0.43, 0.72]	0.88 [0.80, 0.94]
Souza 2014	39	2	10	51	0.80 [0.66, 0.90]	0.96 [0.87, 1.00]
Souza 2018	59	45	21	70	0.74 [0.63, 0.83]	0.61 [0.51, 0.70]
Teo 2014	40	10	6	8	0.87 [0.74, 0.95]	0.44 [0.22, 0.69]



Test 7. VistaProof

VistaProof

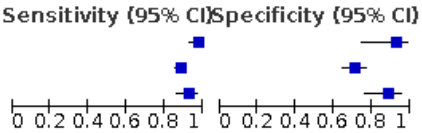
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Achilleos 2013	36	2	0	0	1.00 [0.90, 1.00]	0.00 [0.00, 0.84]
Diniz 2011	38	1	11	5	0.78 [0.63, 0.88]	0.83 [0.36, 1.00]
Diniz 2012	74	1	26	4	0.74 [0.64, 0.82]	0.80 [0.28, 0.99]
Jablonski-Momeni 2011	69	19	4	6	0.95 [0.87, 0.98]	0.24 [0.09, 0.45]
Jablonski-Momeni 2012	57	17	2	6	0.97 [0.88, 1.00]	0.26 [0.10, 0.48]
Jablonski-Momeni 2012a	67	12	0	1	1.00 [0.95, 1.00]	0.08 [0.00, 0.36]
Jablonski-Momeni 2014	50	157	3	96	0.94 [0.84, 0.99]	0.38 [0.32, 0.44]
Jablonski-Momeni 2016	27	11	9	158	0.75 [0.58, 0.88]	0.93 [0.89, 0.97]
Jablonski-Momeni 2016	67	21	8	118	0.89 [0.80, 0.95]	0.85 [0.78, 0.90]
Matos 2011	155	3	195	29	0.44 [0.39, 0.50]	0.91 [0.75, 0.98]
Novaes 2012a	46	12	18	37	0.72 [0.59, 0.82]	0.76 [0.61, 0.87]
Novaes 2016	59	6	17	27	0.78 [0.67, 0.86]	0.82 [0.65, 0.93]
Presoto 2017	37	29	1	40	0.97 [0.86, 1.00]	0.58 [0.45, 0.70]
Rodrigues 2008	76	1	35	7	0.68 [0.59, 0.77]	0.88 [0.47, 1.00]
Rodrigues 2011	25	1	55	16	0.31 [0.21, 0.43]	0.94 [0.71, 1.00]
Seremidi 2012	81	21	2	3	0.98 [0.92, 1.00]	0.13 [0.03, 0.32]
Souza 2013	44	5	16	14	0.73 [0.60, 0.84]	0.74 [0.49, 0.91]
Tonkaboni 2018	25	0	21	62	0.54 [0.39, 0.69]	1.00 [0.94, 1.00]



Test 8. SoproLife

SoproLife

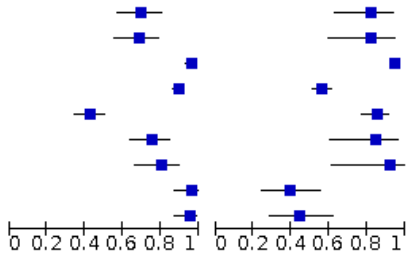
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Kockanat 2017	92	2	2	24	0.98 [0.93, 1.00]	0.92 [0.75, 0.99]
Muller-Bolla 2017	473	60	62	148	0.88 [0.85, 0.91]	0.71 [0.64, 0.77]
Zeitouny 2014	104	6	8	46	0.93 [0.86, 0.97]	0.88 [0.77, 0.96]



Test 9. QLF

QLF

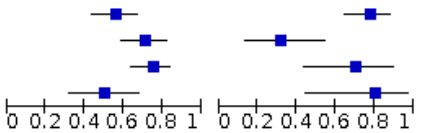
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Bussaneli 2015	46	5	20	23	0.70 [0.57, 0.80]	0.82 [0.63, 0.94]
Diniz 2019	45	4	21	18	0.68 [0.56, 0.79]	0.82 [0.60, 0.95]
Feng 2005	342	75	15	1300	0.96 [0.93, 0.98]	0.95 [0.93, 0.96]
Jung 2018	333	183	39	236	0.90 [0.86, 0.92]	0.56 [0.51, 0.61]
Kim 2017	72	16	98	94	0.42 [0.35, 0.50]	0.85 [0.77, 0.91]
Ko 2015	57	3	19	16	0.75 [0.64, 0.84]	0.84 [0.60, 0.97]
Lee 2018	40	1	10	11	0.80 [0.66, 0.90]	0.92 [0.62, 1.00]
Pereira 2011	53	25	2	16	0.96 [0.87, 1.00]	0.39 [0.24, 0.55]
Yoon 2017	61	21	3	17	0.95 [0.87, 0.99]	0.45 [0.29, 0.62]



Test 10. MidWest

MidWest

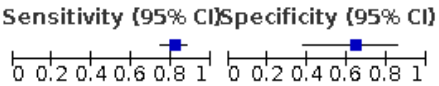
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Aktan 2012	42	12	33	42	0.56 [0.44, 0.67]	0.78 [0.64, 0.88]
Diniz 2019	47	15	19	7	0.71 [0.59, 0.82]	0.32 [0.14, 0.55]
Rodrigues 2011	60	5	20	12	0.75 [0.64, 0.84]	0.71 [0.44, 0.90]
Van Hilsen 2013	16	2	16	8	0.50 [0.32, 0.68]	0.80 [0.44, 0.97]



Test 11. Combined visual/radiograph/DIAGNOdent

Combined visual/radiograph/DIAGNOdent

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Alomari 2015	117	6	26	11	0.82 [0.75, 0.88]	0.65 [0.38, 0.86]



ADDITIONAL TABLES

Table 1. Classification of levels of caries

DMFT classification	Definition (Pitts 2001)
0	Sound (non-diseased)
D_1	Non-cavitated yet clinically detectable enamel lesions with intact surfaces

Table 1. Classification of levels of caries *(Continued)*

D_2	Cavitated lesion penetrating the enamel or shadowing
D_3	Cavity progressing past the enamel-dentine junction into dentine
D_4	Cavity progressing into pulp

DMFT = decayed, missing, and filled teeth.

Table 2. QUADAS-2 tool

Item	Response (delete as required)
Participant selection – Risk of bias	
1) Was a consecutive or random sample of participants or teeth used?	**Yes** – where teeth or participants were selected consecutively or allocated to the study via a randomisation process **No** – if study described another method of sampling **Unclear** – if participant sampling is not described
2) Was a case-control design avoided?	**Yes** – if case-control clearly not used **No** – if study described as case-control or describes sampling specific numbers of participants with particular diagnoses **Unclear** – if not clearly described
3) Did the study avoid inappropriate exclusions (e.g. inclusion of caries into dentine)?	**Yes** – if the study clearly reports that included participants or teeth were apparently healthy or caries into dentine were excluded **No** – if lesions were included that showed caries into dentine or exclusions that might affect test accuracy (e.g. teeth with no caries) **Unclear** – if not clearly reported
Could the selection of participants have introduced bias?	
If answers to all of questions 1) and 2) and 3) was 'yes'	**Risk is low**
If answers to any of questions 1) and 2) and 3) was 'no'	**Risk is high**
If answers to any of questions 1) and 2) and 3) was 'unclear'	**Risk is unclear**
Participant selection – Concerns regarding applicability	
1) Does the study report results for participants or teeth selected by apparent health or suspected early caries (i.e. studies do not recruit patients who are known to have advanced caries into dentine)?	**Yes** – if a group of participants or teeth has been included which is apparently healthy or indicative of early caries **No** – if a group of participants or teeth has been included which is suspected of advanced caries **Unclear** – if insufficient details are provided to determine the spectrum of participants or teeth

Table 2. QUADAS-2 tool *(Continued)*

2) Did the study report data on a per-patient rather than on a tooth or surface basis?	**Yes** – if the analysis was reported on a surface or tooth basis **No** – if the analysis was reported on a per-patient basis **Unclear** - if it is not possible to assess whether data are presented on a per-patient or per-tooth basis
3) Did the study avoid an in vitro setting which required the usage of extracted teeth?	**Yes** – if the participants were recruited prior to tooth extraction **No** – if previously extracted teeth were used in the analysis **Unclear** – if it was not possible to assess the source and method of recruiting of included participants/teeth
Is there concern that the included participants or teeth do not match the review question?	
If answers to all of questions 1) and 2) and 3) was 'yes'	**Risk is low**
If answers to any of questions 1) and 2) and 3) was 'no'	**Risk is high**
If answers to any of questions 1) and 2) and 3) was 'unclear'	**Risk is unclear**
Index test - Risk of bias (to be completed per test evaluated)	
1) Was the index test result interpreted without knowledge of the results of the reference standard?	**Yes** – if the index test described is always conducted and interpreted prior to the reference standard result, or for retrospective studies interpreted without prior knowledge of the reference standard **No** – if index test described as interpreted in knowledge of reference standard result **Unclear** – if index test blinding is not described
2) Was the diagnostic threshold at which the test was considered positive pre-specified?	**Yes** – if threshold was pre-specified (i.e. prior to analysing the study results) **No** – if threshold was not pre-specified **Unclear** – if not possible to tell whether or not diagnostic threshold was pre-specified
For visual and radiograph tests only: 3) For studies reporting the accuracy of multiple diagnostic thresholds for the same index test or multiple index tests, was each threshold or index test interpreted without knowledge of the results of the others?	**Yes** – if thresholds or index tests were selected prospectively and each was interpreted by a different clinician or interpreter, or if study implements a retrospective (or no) cut-off (i.e. look for deepest/most severe lesion first) **No** – if study states reported by same reader **Unclear** - if no mention of number of readers for each threshold or if pre-specification of threshold not reported **N/A** - multiple diagnostic thresholds not reported for the same index test
Could the conduct or interpretation of the index test have introduced bias?	
For visual and radiographic studies item 3) to be added	
If answers to all of questions 1) and 2) was 'yes'	**Risk is low**

Table 2. QUADAS-2 tool *(Continued)*

If answers to any of questions 1) and 2) was 'no'	**Risk is high**
If answers to any of questions 1) and 2) was 'unclear'	**Risk is unclear**
Index test - Concerns regarding applicability	
1) Were thresholds or criteria for diagnosis reported in sufficient detail to allow replication?	**Yes** – if the criteria for detection or diagnosis of the target disorder were reported in sufficient detail to allow replication **No** – if the criteria for detection or diagnosis of the target disorder were not reported in sufficient detail to allow replication **Unclear** - if some but not sufficient information on criteria for diagnosis to allow replication were provided
2) Was the test interpretation carried out by an experienced examiner?	**Yes** – if the test clearly reported that the test was interpreted by an experienced examiner **No** – if the test was not interpreted by an experienced examiner **Unclear** – if the experience of the examiner(s) was not reported in sufficient detail to judge or if examiners described as 'Expert' with no further detail given
Is there concern that the included participants do not match the review question?	
If the answer to question 1) and 2) was 'yes'	**Concern is low**
If the answer to question 1) and 2) was 'no'	**Concern is high**
If the answer to question 1) and 2) was 'unclear'	**Concern is unclear**
Reference standard - Risk of bias	
1) Is the reference standard likely to correctly classify the target condition?	**Yes** – if all teeth or surfaces underwent a histological or excavation reference standard **No** – if a final diagnosis for any participant or tooth was reached without the histological or excavation reference standards **Unclear** – if the method of final diagnosis was not reported
2) Were the reference standard results interpreted without knowledge of the results of the index test?	**Yes** – if the reference standard examiner was described as blinded to the index test result **No** – if the reference standard examiner was described as having knowledge of the index test result **Unclear** – if blinded reference standard interpretation was not clearly reported
Could the reference standard, its conduct, or its interpretation have introduced bias?	
If answers to questions 1) and 2) was 'yes'	**Risk is low**
If the answer to question 1) and 2) was 'no'	**Concern is high**

Table 2. QUADAS-2 tool *(Continued)*

If the answer to question 1) and 2) was 'unclear'	**Concern is unclear**
Reference standard - Concerns regarding applicability	
1) Does the study use the same definition of disease positive as the prescribed in the review question?	**Yes** - same definition of disease positive used, or teeth can be disaggregated and regrouped according to review definition **No** - some teeth cannot be disaggregated **Unclear** - definition of disease positive not clearly reported
Flow and timing - Risk of bias	
1) Was there an appropriate interval between index test and reference standard (in vivo studies less than 3 months, in vitro no limit but must be stored appropriately)?	**Yes** - if study reports index and reference standard had a suitable interval or storage method **No** - if study reports greater than 3-month interval between index and reference standard or inappropriate storage of extracted teeth prior to reference standard **Unclear** - if study does not report interval or storage methods between index and histological reference standard
2) Did all participants receive the same reference standard?	**Yes** - if all participants underwent the same reference standard **No** - if more than 1 reference standard was used **Unclear** - if not clearly reported
3) Were all participants included in the analysis?	**Yes** - if all participants were included in the analysis **No** - if some participants were excluded from the analysis **Unclear** - if not clearly reported
If answers to questions 1) and 2) and 3) was 'yes'	**Risk is low**
If answers to any one of questions 1) or 2) or 3) was 'no'	**Risk is high**
If answers to any one of questions 1) or 2) or 3) was 'unclear'	**Risk is unclear**

N/A = not applicable; QUADAS-2 = Quality Assessment of Diagnostic Accuracy Studies 2.

Table 3. Included studies characteristics

Study ID	Test	Number of sites reported	Number of teeth included	Number of participants	In vitro/in vivo	Threshold	Prevalence of enamel caries	Prevalence of dentine caries	Surface	Reference standard	Dentition
Achilleos 2013	DIAGNOdent pen	38	38	NR - extracted	vitro	13	0.95	0.39	Occlusal	Histology	Permanent
Achilleos 2013	VistaProof	38	38	NR - extracted	vitro	1	0.95	0.39	Occlusal	Histology	Permanent
Akarsu 2006	DIAGNOdent	165	187	161	vivo	5.5	0.77	0.52	Occlusal	Excavation	Permanent
Aktan 2012	DIAGNOdent pen	129	83	NR – extracted	vitro	13	0.58	0.21	Occlusal	Histology	Permanent
Aktan 2012	MidWest	129	83	NR - extracted	vitro	N/A	0.58	0.21	Occlusal	Histology	Permanent
Almosa 2014	DIAGNOdent pen	1653	822	89	vivo	13	0.33	0.01	Smooth	Visual	Permanent
Alomari 2015	Combined visual/radiograph/DIAGNOdent	160	NR	NR - extracted	vitro	N/A	0.89	0.38	Occlusal	Histology	Permanent
Apostolopoulou 2009	DIAGNOdent	111	24	NR - extracted	vitro	NR	0.98	0.22	Occlusal	Histology	Primary
Attrill 2001	DIAGNOdent	58	58	NR - extracted	vitro	9	0.60	0.51	Occlusal	Histology	Primary
Bahrololoomi 2015	DIAGNOdent	109	115	31	vivo	8	0.94	0.52	Occlusal	Excavation	Permanent
Bamzahim 2004	DIAGNOdent	66	66	NR - extracted	vitro	10	0.52	NR	Secondary	Histology	Permanent
Baseren 2003	DIAGNOdent	31	35	NR - extracted	vitro	13	0.39	0.19	Occlusal	Histology	Permanent
Bittar 2012	DIAGNOdent pen	55	33	NR - extracted	vitro	8	0.67	0.22	Occlusal	Histology	Primary
Bittar 2012a	DIAGNOdent pen	58	33	NR - extracted	vitro	7	0.62	0.28	Approximal	Histology	Primary
Braga 2009	DIAGNOdent pen	131	84	NR - extracted	vitro	4	0.63	0.26	Occlusal	Histology	Primary

Table 3. Included studies characteristics *(Continued)*

Bussaneli 2015	DIAGNOdent pen	94	102	NR - extracted	vitro	15	0.70	0.19	Occlusal	Histology	Permanent
Bussaneli 2015	QLF	94	102	NR - extracted	vitro	-10.5	0.70	0.19	Occlusal	Histology	Permanent
Bussaneli 2015a	DIAGNOdent pen	59	59	45	vitro	15	0.71	0.58	Approximal	Visual	Primary
Castilho 2016	DIAGNOdent	43	43	26	vivo	6	0.81	0.07	Occlusal	Histology	Permanent
Chen 2012	DIAGNOdent	256	216	96	vivo	7	0.50	0.35	Approximal	Excavation	Primary
Chong 2003	DIAGNOdent	320	NR	NR - extracted	vitro	5	0.15	0.06	Occlusal	Visual	Permanent
Cinar 2013	DIAGNOdent	44	NR	NR - extracted	vitro	6	0.75	0.2	Occlusal	Histology	Primary
Cinar 2013	DIAGNOdent pen	44	NR	NR - extracted	vitro	14	0.75	0.2	Occlusal	Histology	Primary
Costa 2002	DIAGNOdent	49	49	NR - extracted	vitro	6	0.65	0.31	Occlusal	Histology	Permanent
Diniz 2011	DIAGNOdent	55	55	NR - extracted	vitro	16	0.89	0.11	Occlusal	Histology	Permanent
Diniz 2011	DIAGNOdent pen	55	55	NR - extracted	vitro	11	0.89	0.11	Occlusal	Histology	Permanent
Diniz 2011	VistaProof	55	55	NR - extracted	vitro	1.2	0.89	0.11	Occlusal	Histology	Permanent
Diniz 2012	DIAGNOdent	105	105	88	vitro	16	0.95	0.28	Occlusal	Histology	Permanent
Diniz 2012	DIAGNOdent pen	105	105	88	vitro	11	0.95	0.28	Occlusal	Histology	Permanent
Diniz 2012	VistaProof	105	105	88	vitro	1	0.95	0.28	Occlusal	Histology	Permanent
Diniz 2019	DIAGNOdent	88	88	NR - extracted	vitro	5	0.75	0.63	Occlusal	Histology	Primary
Diniz 2019	DIAGNOdent pen	87	88	NR - extracted	vitro	4	0.76	0.63	Occlusal	Histology	Primary
Diniz 2019	QLF	88	88	NR - extracted	vitro	7.5	0.75	0.63	Occlusal	Histology	Primary
Diniz 2019	MidWest	88	88	NR - extracted	vitro	N/A	0.75	0.63	Occlusal	Histology	Primary
Duruturk 2011	DIAGNOdent	505	505	307	vivo	15	0.36	0.36	Occlusal	Visual	Primary
Feng 2005	QLF	1732	1732	300	vivo	Exam-iner	0.21		Occlusal	Visual	Permanent

Table 3. Included studies characteristics *(Continued)*

Goel 2009	DIAGNOdent	83	84	NR	vivo	6	0.98	0.43	Occlusal	Histology	Permanent
Hein-rich-Weltzien 2003	DIAGNOdent	248	248	94	vivo	NR	0.90	0.85	Occlusal	Excavation	Permanent
Huth 2010	DIAGNOdent pen	117	117	117	vivo	7	0.66	0.37	Occlusal	Excavation	Permanent
Iranzo-Cortes 2017	DIAGNOdent	64	65	NR - extracted	vitro	14	0.77	0.17	Occlusal	Histology	Permanent
Jablonski-Mo-meni 2011	VistaProof	98	53	NR - extracted	vitro	0.9	0.74	0.23	Occlusal	Histology	Permanent
Jablonski-Mo-meni 2012	DIAGNOdent pen	82	36	NR - extracted	vitro	6	0.72	0.21	Occlusal	Histology	Permanent
Jablonski-Mo-meni 2012	VistaProof	82	36	NR - extracted	vitro	0.9	0.72	0.21	Occlusal	Histology	Permanent
Jablonski-Mo-meni 2012a	DIAGNOdent	84	36	NR - extracted	vitro	8	0.85	0.48	Occlusal	Histology	Permanent
Jablonski-Mo-meni 2012a	VistaProof	80	36	NR - extracted	vitro	0.9	0.84	0.48	Occlusal	Histology	Permanent
Jablonski-Mo-meni 2014	VistaProof	306	26	NR	vivo	1	0.17	0.12	Occlusal	Visual	Permanent
Jablonski-Mo-meni 2016	VistaProof	205	205	35	vivo	1.3	0.18	0	Occlusal	Visual	Primary
Jablonski-Mo-meni 2016a	VistaProof	214	214	35	vivo	1.3	0.35	0	Occlusal	Visual	Permanent
Jung 2018	QLF	791	791	94	vitro	0.47	0.47	0.14	Occlusal	Visual	Permanent
Kim 2017	QLF	280	280	65	vitro	Exam-iner	0.61	0.2	Approximal	Radiograph	Permanent
Ko 2015	QLF	95	120	NR	vivo	11	0.80	0.15	Approximal	Histology	Permanent
Kockanat 2017	DIAGNOdent pen	120	144	NR	vivo	14	0.78	0.32	Occlusal	Histology	Primary

Table 3. Included studies characteristics *(Continued)*

Kockanat 2017	SoproLife	120	144	NR	vivo	Exam-iner	0.78	0.32	Occlusal	Histology	Primary
Kouchaji 2012	DIAGNOdent	156	156	40	vivo	15	0.85	0.29	Occlusal	Visual	Permanent
Kucukyilmaz 2015	DIAGNOdent	200	200	200	vivo	15	0.82	0.33	Occlusal	Histology	Primary
Kuhnisch 2008	DIAGNOdent	840	840	311	vivo	16	0.71	0.06	Occlusal	Visual	Mixed
Lee 2018	QLF	62	66	NR - extracted	vitro	NR	0.81	0.11	Occlusal	Histology	Permanent
Lussi 2006	DIAGNOdent pen	150	150	75	vitro	6	0.59	0.25	Approximal	Histology	Permanent
Lussi 2006a	DIAGNOdent	119	119	NR - extracted	vitro	7	0.78	0.35	Occlusal	Histology	Primary
Lussi 2006a	DIAGNOdent pen	119	119	NR - extracted	vitro	6	0.78	0.35	Occlusal	Histology	Primary
Mansour 2016	DIAGNOdent	426	932	40	vitro	14	0.12	0.14	Occlusal	Visual	Permanent
Matos 2011	DIAGNOdent pen	382	382	68	vivo	4	0.92	0.05	Occlusal	Visual	Primary
Matos 2011	VistaProof	382	382	68	vivo	1.1	0.92	0.05	Occlusal	Visual	Primary
Mendes 2005	DIAGNOdent	77	77	NR - extracted	vitro	4	0.69	0.14	Smooth	Histology	Primary
Mendes 2006	DIAGNOdent	110	79	NR - extracted	vitro	7	0.75	0.25	Occlusal	Histology	Primary
Mepparambath 2014	DIAGNOdent	169	101	NR	vivo	10	0.22	0.08	Approximal	Radiograph	Primary
Mortensen 2018	DIAGNOdent pen	60	60	57	vivo	12	0.97	0.45	Occlusal	Visual	Permanent
Muller-Bolla 2017	DIAGNOdent pen	743	743	103	vivo	12	0.72	0.29	Occlusal	Visual	Mixed
Muller-Bolla 2017	SoproLife	743	743	103	vivo	N/A	0.72	0.29	Occlusal	Visual	Mixed
Neuhaus 2011	DIAGNOdent	37	37	NR - extracted	vitro	9	0.73	0.24	Occlusal	Histology	Primary
Novaes 2009	DIAGNOdent pen	621	50	NR	vivo	5	0.41	0.03	Approximal	Visual	Primary

Table 3. Included studies characteristics *(Continued)*

Novaes 2010	DIAGNOdent pen	592	168	76	vivo	5	0.81	0.05	Approximal	Visual	Primary
Novaes 2012a	DIAGNOdent	113	113	77	vitro	7	0.57	0.17	Occlusal	Histology	Primary
Novaes 2012a	DIAGNOdent pen	113	113	77	vitro	8	0.57	0.17	Occlusal	Histology	Primary
Novaes 2012a	VistaProof	113	113	77	vitro	Examiner	0.57	0.17	Occlusal	Histology	Primary
Novaes 2016	DIAGNOdent	109	109	65	vitro	2	0.70	0.23	Smooth	Histology	Primary
Novaes 2016	DIAGNOdent pen	109	109	65	vitro	3	0.70	0.23	Smooth	Histology	Primary
Novaes 2016	VistaProof	109	109	65	vitro	1.1	0.70	0.23	Smooth	Histology	Primary
Ozsevik 2015	DIAGNOdent pen	156	156	87	vitro	9	0.63	0.35	Approximal	Histology	Permanent
Paula 2011	DIAGNOdent	64	64	26	vitro	10	0.88	0.28	Occlusal	Histology	Permanent
Pereira 2011	DIAGNOdent	96	96	NR - extracted	vitro	5	0.57	0.25	Occlusal	Histology	Permanent
Pereira 2011	QLF	96	96	NR - extracted	vitro	Examiner	0.57	0.25	Occlusal	Histology	Permanent
Pinelli 2002	DIAGNOdent	220	220	50	vivo	4	0.50	NR	Smooth	Visual	Permanent
Presoto 2017	VistaProof	107	107	14	vivo	N/A	0.36	NR	Occlusal	Visual	Permanent
Randomeirelles 2011	DIAGNOdent	789	789	179	vivo	20	0.34	0.31	Occlusal	Radiograph	Mixed
Ribeiro 2015	DIAGNOdent pen	63	137	112	vivo	6	0.60	0.29	Approximal	Radiograph	Primary
Rocha 2003	DIAGNOdent	100	50	29	vivo	6	0.58	0.14	Occlusal	Histology	Primary
Rodrigues 2008	DIAGNOdent	119	119	NR - extracted	vitro	7	0.65	0.54	Occlusal	Histology	Permanent
Rodrigues 2008	DIAGNOdent pen	119	119	NR - extracted	vitro	6	0.65	0.54	Occlusal	Histology	Permanent
Rodrigues 2008	VistaProof	119	119	NR - extracted	vitro	1.26	0.65	0.54	Occlusal	Histology	Permanent

Table 3. Included studies characteristics *(Continued)*

Rodrigues 2009	DIAGNOdent	169	169	NR - extracted	vitro	7	0.93	0.11	Occlusal	Histology	Permanent
Rodrigues 2009a	DIAGNOdent	148	148	NR - extracted	vitro	7	0.92	0.03	Occlusal	Histology	Primary
Rodrigues 2011	DIAGNOdent	97	97	NR - extracted	vitro	NR	0.82	0.28	Occlusal	Histology	Permanent
Rodrigues 2011	DIAGNOdent pen	97	97	NR - extracted	vitro	NR	0.82	0.28	Occlusal	Histology	Permanent
Rodrigues 2011	VistaProof	97	97	NR - extracted	vitro	NR	0.82	0.28	Occlusal	Histology	Permanent
Rodrigues 2011	MidWest	97	97	NR - extracted	vitro	N/A	0.82	0.28	Occlusal	Histology	Permanent
Seremidi 2012	DIAGNOdent pen	107	107	41	vitro	9	0.78	0.19	Occlusal	Histology	Permanent
Seremidi 2012	VistaProof	107	107	41	vitro	1.3	0.78	0.19	Occlusal	Histology	Permanent
Sheehy 2001	DIAGNOdent	170	107	41	vitro	14	0.55	0.28	Occlusal	Visual	Permanent
Shi 2000	DIAGNOdent	70	76	NR - extracted	vitro	NR	0.73	0.39	Occlusal	Histology	Permanent
Souza 2013	DIAGNOdent	79	79	NR - extracted	vitro	15	0.76	0.35	Occlusal	Histology	Permanent
Souza 2013	DIAGNOdent pen	79	79	NR - extracted	vitro	19	0.76	0.35	Occlusal	Histology	Permanent
Souza 2013	VistaProof	79	79	NR - extracted	vitro	1.1	0.76	0.35	Occlusal	Histology	Permanent
Souza 2014	DIAGNOdent pen	102	102	NR - extracted	vitro	27	0.48	0.34	Approximal	Histology	Permanent
Souza 2014a	DIAGNOdent pen	144	144	72	vitro	27	0.35	0.1	Approximal	Histology	Primary
Souza 2018	DIAGNOdent pen	195	195	46	vitro	13	0.41	0.13	Approximal	Visual	Primary
Sridhar 2009	DIAGNOdent	50	50	NR - extracted	vitro	5	0.88	0.12	Occlusal	Histology	Permanent
Teo 2014	DIAGNOdent pen	64	64	NR - extracted	vitro	9	0.72	0.31	Occlusal	Histology	Permanent
Tonkaboni 2018	VistaProof	108	108	NR - extracted	vitro	Examiner	0.43	0.35	Approximal	Histology	Permanent
Umemori 2010	DIAGNOdent	100	100	19	vitro	NR	0.36	0.12	Occlusal	Visual	Permanent

Table 3. Included studies characteristics *(Continued)*

Van Hilsen 2013	MidWest	42	45	NR - extracted	vitro	N/A	0.76	0.31	Occlusal	Histology	Permanent
Virajsilp 2005	DIAGNOdent	107	72	NR - extracted	vitro	NR	0.83	0.5	Approximal	Histology	Primary
Yoon 2017	DIAGNOdent	102	102	NR - extracted	vitro	10	0.63	NR	Approximal	Radiograph	Permanent
Yoon 2017	QLF	102	102	NR - extracted	vitro	-13.8	0.63	NR	Approximal	Radiograph	Permanent
Zeitouny 2014	SoproLife	164	219	NR - extracted	vitro	Exam-iner	0.68	0.68	Occlusal	Visual	Permanent

N/A = not applicable; NR = not reported; QLF = quantitative light-induced fluorescence.

Table 4. Investigations of test type, dentition, and potential sources of heterogeneity in all studies

Test	Datasets	Tooth surfaces (caries)	DOR (95% CI)	RDOR (95% CI)	P value (LR)
Difference between blue, green, and red fluorescence					
Blue	21	3429 (2163)	18.47 (10.59 to 32.20)	1.0 (comparator)	0.14
Green	9	3340 (1276)	19.49 (9.01 to 42.18)	1.06 (0.41 to 2.73)	
Red	84	14,514 (8705)	12.75 (9.74 to 16.68)	0.69 (0.37 to 1.28)	
Difference between permanent/mixed and primary dentition					
Permanent/mixed	74	13,427 (7195)	15.21 (11.35 to 20.37)	1.0 (comparator)	0.19
Primary	40	6024 (3885)	12.34 (8.44 to 18.04)	0.81 (0.50 to 1.31)	
Difference between prevalence of dentine caries in sample (low 14%, medium 15% to 34%, high ≥ 35%)					
Low	26	7899 (4118)	11.10 (6.88 to 17.91)	0.76 (0.39 to 1.48)	0.32
Medium	57	8868 (5057)	15.39 (11.07 to 21.41)	1.05 (0.59 to 1.86)	
High	31	3688 (2593)	14.59 (9.18 to 23.22)	1.0 (comparator)	
Difference between occlusal, proximal, and smooth surfaces					
Occlusal	89	15,204 (9252)	14.29 (10.92 to 18.72)	1.10 (0.59 to 2.02)	0.62
Proximal	18	3490 (1983)	13.06 (7.52 to 22.67)	1.0 (comparator)	
Smooth	7	2277 (919)	13.41 (5.58 to 32.25)	1.03 (0.36 to 2.90)	
Difference in reference standard					
Histology/excavation	83	7875 (5609)	13.49 (10.25 to 17.76)	1.0 (comparator)	0.06
Visual	25	10,762 (5282)	19.32 (12.37 to 30.17)	1.43 (0.85 to 2.41)	
Radiography	6	1505 (639)	6.21 (2.58 to 14.97)	0.46 (0.18 to 1.16)	
Difference between multiple or single sites per tooth					
Multiple	24	4371 (2999)	9.46 (5.91 to 15.14)	0.59 (0.35 to 1.02)	0.06
Single	90	16,666 (9189)	15.96 (12.26 to 20.77)	1.0 (comparator)	

CI = confidence interval; DOR = diagnostic odds ratio; LR = likelihood ratio; RDOR = relative diagnostic odds ratio.

APPENDICES

Appendix 1. MEDLINE Ovid search strategy

1. exp Tooth demineralization/
2. (teeth adj5 (cavit$ or caries or carious or decay$ or lesion$ or demineral$ or remineral$)).mp.
3. (tooth adj5 (cavit$ or caries or carious or decay$ or lesion$ or demineral$ or remineral$)).mp.
4. (dental adj5 (cavit$ or caries or carious or decay$ or lesion$ or demineral$ or remineral$)).mp.
5. (enamel adj5 (cavit$ or caries or carious or decay$ or lesion$ or demineral$ or remineral$)).mp.
6. (dentin$ adj5 (cavit$ or caries or carious or decay$ or lesion$ or demineral$ or remineral$)).mp.
7. (root adj5 (cavit$ or caries or carious or decay$ or lesion$ or demineral$ or remineral$)).mp.
8. or/1-7
9. Fluorescence/
10. exp Lasers/
11. fluorescen$.mp.
12. (QLF or DiagnoDENT).mp.
13. ((ultraviolet$ or light$ or laser$) adj5 (detect$ or diagnos$)).mp.
14. (quantitative adj (light$ or laser$)).mp.
15. or/9-14
16. 8 and 15

Appendix 2. Embase Ovid search strategy

1. dental caries/
2. (teeth adj5 (cavit$ or caries or carious or decay$ or lesion$ or demineral$ or remineral$)).mp.
3. (tooth adj5 (cavit$ or caries or carious or decay$ or lesion$ or demineral$ or remineral$)).mp.
4. (dental adj5 (cavit$ or caries or carious or decay$ or lesion$ or demineral$ or remineral$)).mp.
5. (enamel adj5 (cavit$ or caries or carious or decay$ or lesion$ or demineral$ or remineral$)).mp.
6. (dentin$ adj5 (cavit$ or caries or carious or decay$ or lesion$ or demineral$ or remineral$)).mp.
7. (root adj5 (cavit$ or caries or carious or decay$ or lesion$ or demineral$ or remineral$)).mp.
8. or/1-7
9. Fluorescence/
10. exp Lasers/
11. fluorescen$.mp.
12. (QLF or DiagnoDENT).mp.
13. ((ultraviolet$ or light$ or laser$) adj5 (detect$ or diagnos$)).mp.
14. (quantitative adj (light$ or laser$)).mp.
15. or/9-14
16. 8 and 15

Appendix 3. US National Institutes of Health Ongoing Trials Register (ClinicalTrials.gov) search strategy

Expert search interface: (caries OR tooth decay OR dental decay OR cavities OR carious) AND (fluorescence OR QLF OR laser OR DiagnoDENT OR ultraviolet OR light) AND (diagnosis OR diagnose OR detect OR detection)

Appendix 4. World Health Organization International Clinical Trials Registry Platform search strategy

caries AND fluorescence OR caries AND QLF OR caries AND DiagnoDENT OR caries AND ultraviolet OR caries AND light

caries AND laser AND diagnosis OR caries AND laser AND detection

Appendix 5. Comparison of fluorescence devices

Test	Studies	Teeth (caries)	DOR (95% CI)	RDOR (95% CI)	P value (LR)
Difference between red fluorescence studies					
DIAGNOdent	46	7316 (4363)	16.03 (11.14 to 23.05)	1.40 (0.93 to 2.12)	0.71
DIAGNOdent pen	34	6842 (4089)	11.44 (8.12 to 16.11)		

(Continued)

MidWest[a]	4	356 (253)	39.39 (2.44 to 635.99)	-	-
Difference between blue fluorescence studies					
SoproLife	3	1027 (741)	69.75 (24.32 to 200.01)	4.75 (1.46 to 15.45)	0.0095
VistaProof and Cam	18	2402 (1422)	14.66 (8.58 to 25.04)		
Difference between green fluorescence studies					
QLF software-based decision	7	2964 (1051)	24.10 (8.60 to 67.90)	3.10 (0.38 to 25.07)	0.34
QLF image-based decision	2	376 (225)	8.20 (2.40 to 28.10)		

[a]MidWest not included in analysis due to small number of studies and low DOR.

CI = confidence interval; DOR = diagnostic odds ratio; LR = likelihood ratio; QLF = quantitative light-induced fluorescence; RDOR = relative diagnostic odds ratio.

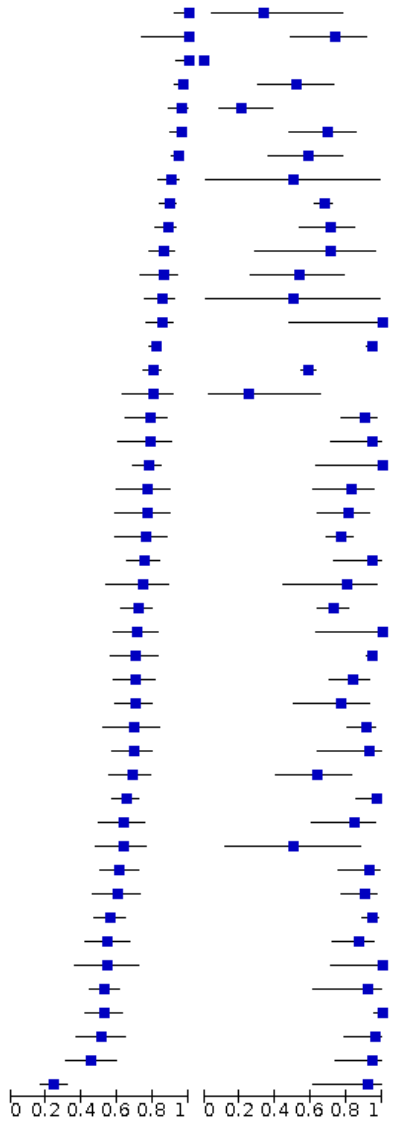
Red fluorescence

We included 84 datasets that used a laser fluorescence device. This included 46 DIAGNOdent, 34 DIAGNOdent pen and four MidWest, which together assessed 14,514 tooth surfaces. The Canary System was not used by any included study. The findings of individual studies subgrouped by the device used are shown in the forest plots in Figure 22 and the hierarchical summary receiver operating characteristic (HSROC) curves for each group of devices are plotted in Figure 23. 10 studies investigated DIAGNOdent and DIAGNOdent pen (Cinar 2013; Diniz 2011; Diniz 2012; Diniz 2019; Lussi 2006a; Novaes 2012a; Novaes 2016; Rodrigues 2008; Rodrigues 2011; Souza 2013) and three studies compared MidWest to DIAGNOdent pen (Aktan 2012; Diniz 2019; Rodrigues 2011).

Figure 22. Forest plot of tests investigating laser fluorescence devices: DIAGNOdent, DIAGNOdent pen, and MidWest.

DIAGNOdent

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Sridhar 2009	44	4	0	2	1.00 [0.92, 1.00]	0.33 [0.04, 0.78]
Baseren 2003	12	5	0	14	1.00 [0.74, 1.00]	0.74 [0.49, 0.91]
Chong 2003	48	272	0	0	1.00 [0.93, 1.00]	0.00 [0.00, 0.01]
Kouchaji 2012	129	11	4	12	0.97 [0.92, 0.99]	0.52 [0.31, 0.73]
Novaes 2016	73	26	3	7	0.96 [0.89, 0.99]	0.21 [0.09, 0.39]
Lussi 2006a	89	8	4	18	0.96 [0.89, 0.99]	0.69 [0.48, 0.86]
Heinrich-Weltzien 2003	212	10	12	14	0.95 [0.91, 0.97]	0.58 [0.37, 0.78]
Apostolopoulou 2009	98	1	11	1	0.90 [0.83, 0.95]	0.50 [0.01, 0.99]
Duruturk 2011	163	105	20	217	0.89 [0.84, 0.93]	0.67 [0.62, 0.72]
Akarsu 2006	112	11	15	27	0.88 [0.81, 0.93]	0.71 [0.54, 0.85]
Bahrololoomi 2015	88	2	14	5	0.86 [0.78, 0.92]	0.71 [0.29, 0.96]
Iranzo-Cortes 2017	42	7	7	8	0.86 [0.73, 0.94]	0.53 [0.27, 0.79]
Goel 2009	69	1	12	1	0.85 [0.76, 0.92]	0.50 [0.01, 0.99]
Diniz 2012	85	0	15	5	0.85 [0.76, 0.91]	1.00 [0.48, 1.00]
Kuhnisch 2008	489	13	110	228	0.82 [0.78, 0.85]	0.95 [0.91, 0.97]
Rando-Meirelles 2011	213	215	53	308	0.80 [0.75, 0.85]	0.59 [0.55, 0.63]
Castilho 2016	28	6	7	2	0.80 [0.63, 0.92]	0.25 [0.03, 0.65]
Pereira 2011	43	4	12	37	0.78 [0.65, 0.88]	0.90 [0.77, 0.97]
Costa 2002	25	1	7	16	0.78 [0.60, 0.91]	0.94 [0.71, 1.00]
Rodrigues 2008	86	0	25	8	0.77 [0.69, 0.85]	1.00 [0.63, 1.00]
Attrill 2001	27	4	8	19	0.77 [0.60, 0.90]	0.83 [0.61, 0.95]
Bamzahim 2004	26	6	8	26	0.76 [0.59, 0.89]	0.81 [0.64, 0.93]
Mepparambath 2014	28	31	9	101	0.76 [0.59, 0.88]	0.77 [0.68, 0.83]
Virajsilp 2005	67	1	22	17	0.75 [0.65, 0.84]	0.94 [0.73, 1.00]
Neuhaus 2011	20	2	7	8	0.74 [0.54, 0.89]	0.80 [0.44, 0.97]
Pinelli 2002	78	30	31	81	0.72 [0.62, 0.80]	0.73 [0.64, 0.81]
Paula 2011	40	0	16	8	0.71 [0.58, 0.83]	1.00 [0.63, 1.00]
Mansour 2016	36	21	15	354	0.71 [0.56, 0.83]	0.94 [0.92, 0.97]
Novaes 2012a	45	8	19	41	0.70 [0.58, 0.81]	0.84 [0.70, 0.93]
Rodrigues 2011	56	4	24	13	0.70 [0.59, 0.80]	0.76 [0.50, 0.93]
Umemori 2010	25	6	11	58	0.69 [0.52, 0.84]	0.91 [0.81, 0.96]
Jablonski-Momeni 2012a	49	1	22	12	0.69 [0.57, 0.79]	0.92 [0.64, 1.00]
Diniz 2019	45	8	21	14	0.68 [0.56, 0.79]	0.64 [0.41, 0.83]
Kucukyilmaz 2015	107	1	57	35	0.65 [0.57, 0.72]	0.97 [0.85, 1.00]
Souza 2013	38	3	22	16	0.63 [0.50, 0.75]	0.84 [0.60, 0.97]
Diniz 2011	31	3	18	3	0.63 [0.48, 0.77]	0.50 [0.12, 0.88]
Mendes 2006	51	2	32	25	0.61 [0.50, 0.72]	0.93 [0.76, 0.99]
Rocha 2003	35	4	23	38	0.60 [0.47, 0.73]	0.90 [0.77, 0.97]
Chen 2012	72	7	56	121	0.56 [0.47, 0.65]	0.95 [0.89, 0.98]
Yoon 2017	35	5	29	33	0.55 [0.42, 0.67]	0.87 [0.72, 0.96]
Cinar 2013	18	0	15	11	0.55 [0.36, 0.72]	1.00 [0.72, 1.00]
Rodrigues 2009	83	1	74	11	0.53 [0.45, 0.61]	0.92 [0.62, 1.00]
Sheehy 2001	49	0	44	77	0.53 [0.42, 0.63]	1.00 [0.95, 1.00]
Mendes 2005	27	1	26	23	0.51 [0.37, 0.65]	0.96 [0.79, 1.00]
Shi 2000	23	1	28	18	0.45 [0.31, 0.60]	0.95 [0.74, 1.00]
Rodrigues 2009	33	1	103	11	0.24 [0.17, 0.32]	0.92 [0.62, 1.00]



DIAGNOdent pen

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Bussaneli 2015a	41	6	1	11	0.98 [0.87, 1.00]	0.65 [0.38, 0.86]
Bittar 2012	36	10	1	8	0.97 [0.86, 1.00]	0.44 [0.22, 0.69]
Novaes 2016	68	18	8	15	0.89 [0.80, 0.95]	0.45 [0.28, 0.64]
Kockanat 2017	84	0	10	26	0.89 [0.81, 0.95]	1.00 [0.87, 1.00]
Aktan 2012	67	29	8	25	0.89 [0.80, 0.95]	0.46 [0.33, 0.60]
Rodrigues 2008	99	1	12	7	0.89 [0.82, 0.94]	0.88 [0.47, 1.00]
Diniz 2012	89	1	11	4	0.89 [0.81, 0.94]	0.80 [0.28, 0.99]
Lussi 2006a	82	6	11	20	0.88 [0.80, 0.94]	0.77 [0.56, 0.91]
Lussi 2006	78	5	11	56	0.88 [0.79, 0.94]	0.92 [0.82, 0.97]
Teo 2014	40	10	6	8	0.87 [0.74, 0.95]	0.44 [0.22, 0.69]
Braga 2009	71	37	11	12	0.87 [0.77, 0.93]	0.24 [0.13, 0.39]
Cinar 2013	28	2	5	9	0.85 [0.68, 0.95]	0.82 [0.48, 0.98]
Diniz 2011	41	3	8	3	0.84 [0.70, 0.93]	0.50 [0.12, 0.88]
Jablonski-Momeni 2012	49	8	10	15	0.83 [0.71, 0.92]	0.65 [0.43, 0.84]
Diniz 2019	53	11	13	10	0.80 [0.69, 0.89]	0.48 [0.26, 0.70]
Souza 2014	39	2	10	51	0.80 [0.66, 0.90]	0.96 [0.87, 1.00]

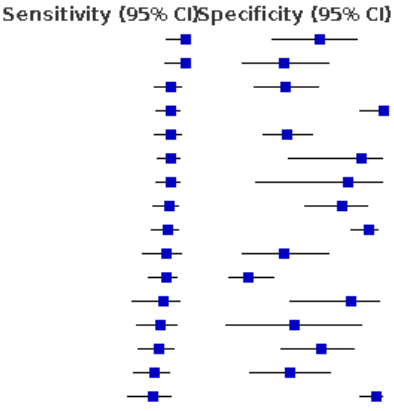
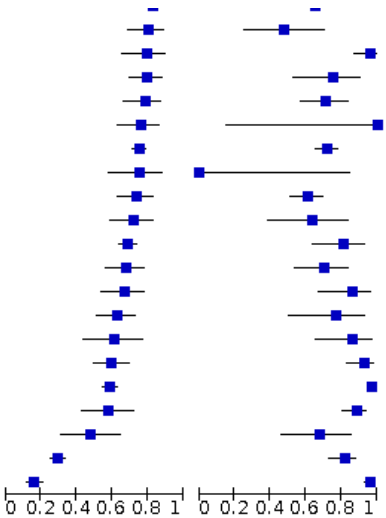


Figure 22. (Continued)

Jabloński-Moreni 2012	49	8	10	15	0.83 [0.71, 0.92]	0.63 [0.43, 0.84]
Diniz 2019	53	11	13	10	0.80 [0.69, 0.89]	0.48 [0.26, 0.70]
Souza 2014	39	2	10	51	0.80 [0.66, 0.90]	0.96 [0.87, 1.00]
Seremidi 2012	66	6	17	18	0.80 [0.69, 0.88]	0.75 [0.53, 0.90]
Novaes 2012a	50	14	14	35	0.78 [0.66, 0.87]	0.71 [0.57, 0.83]
Mortensen 2018	44	0	14	2	0.76 [0.63, 0.86]	1.00 [0.16, 1.00]
Muller-Bolla 2017	404	58	133	148	0.75 [0.71, 0.79]	0.72 [0.65, 0.78]
Achilleos 2013	27	2	9	0	0.75 [0.58, 0.88]	0.00 [0.00, 0.84]
Souza 2018	59	45	21	70	0.74 [0.63, 0.83]	0.61 [0.51, 0.70]
Souza 2013	43	7	17	12	0.72 [0.59, 0.83]	0.63 [0.38, 0.84]
Matos 2011	240	6	110	26	0.69 [0.63, 0.73]	0.81 [0.64, 0.93]
Huth 2010	52	12	25	28	0.68 [0.56, 0.78]	0.70 [0.53, 0.83]
Bussaneli 2015	44	4	22	24	0.67 [0.54, 0.78]	0.86 [0.67, 0.96]
Rodrigues 2011	50	4	30	13	0.63 [0.51, 0.73]	0.76 [0.50, 0.93]
Bittar 2012	22	3	14	19	0.61 [0.43, 0.77]	0.86 [0.65, 0.97]
Ozsevik 2015	59	4	40	53	0.60 [0.49, 0.69]	0.93 [0.83, 0.98]
Almosa 2014	317	34	223	1079	0.59 [0.54, 0.63]	0.97 [0.96, 0.98]
Souza 2014	29	11	21	83	0.58 [0.43, 0.72]	0.88 [0.80, 0.94]
Ribeiro 2015	18	8	20	17	0.47 [0.31, 0.64]	0.68 [0.46, 0.85]
Novaes 2010	141	21	337	93	0.29 [0.25, 0.34]	0.82 [0.73, 0.88]
Novaes 2009	41	15	215	350	0.16 [0.12, 0.21]	0.96 [0.93, 0.98]



MidWest

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Rodrigues 2011	60	5	20	12	0.75 [0.64, 0.84]	0.71 [0.44, 0.90]
Diniz 2019	47	15	19	7	0.71 [0.59, 0.82]	0.32 [0.14, 0.55]
Aktan 2012	42	12	33	42	0.56 [0.44, 0.67]	0.78 [0.64, 0.88]
Van Hilsen 2013	16	2	16	8	0.50 [0.32, 0.68]	0.80 [0.44, 0.97]

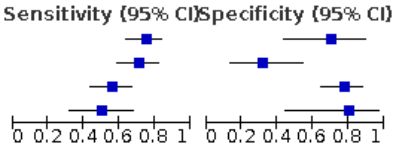
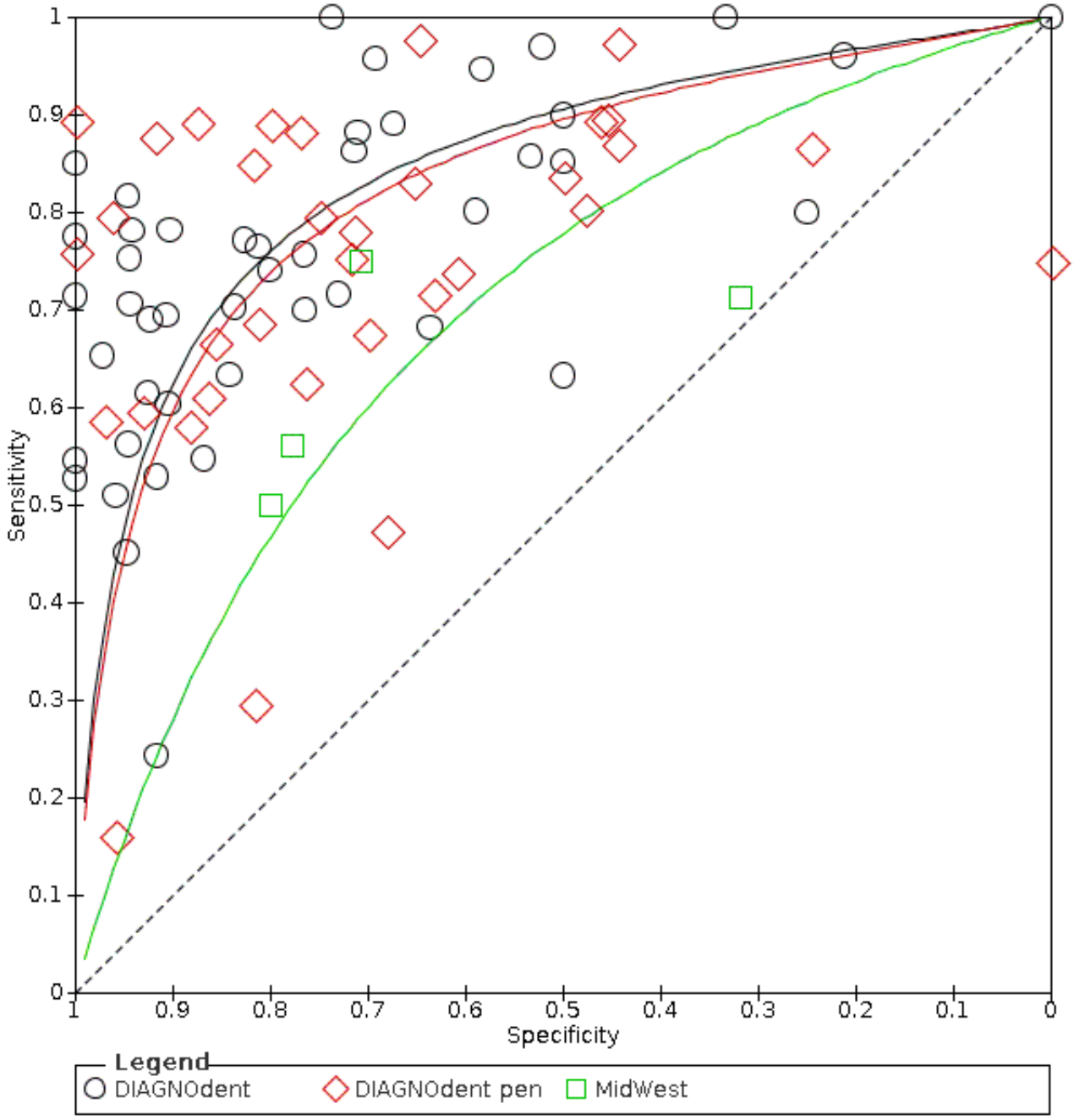


Figure 23. Summary receiver operating characteristic (SROC) plot of fluorescence tests investigating differences between DIAGNOdent, DIAGNOdent pen, and MidWest.



There was no difference observed between the accuracy of the DIAGNOdent and DIAGNOdent pen devices (RDOR 1.40 (95% CI 0.93 to 2.12); P = 0.71).

Blue fluorescence

We found 21 datasets that used blue fluorescence methods to detect caries. This included 18 that investigated the VistaProof device (Achilleos 2013; Diniz 2011; Diniz 2012; Jablonski-Momeni 2011; Jablonski-Momeni 2012; Jablonski-Momeni 2012a; Jablonski-Momeni 2014; Jablonski-Momeni 2016; Matos 2011; Novaes 2012a; Novaes 2016; Presoto 2017; Rodrigues 2008; Rodrigues 2011; Seremidi 2012; Souza 2013; Tonkaboni 2018) and three SoproLife (Kockanat 2017; Muller-Bolla 2017; Zeitouny 2014). The Spectra caries detection device

also fits into this category but no studies provided data for inclusion in the meta-analysis (Markowitz 2015). Individual study estimates of sensitivity and specificity are shown in Figure 24 and summary receiver operating characteristic (SROC) estimates are shown in Figure 25.

Figure 24. Forest plot of tests of red fluorescence devices: VistaProof and SoproLife.

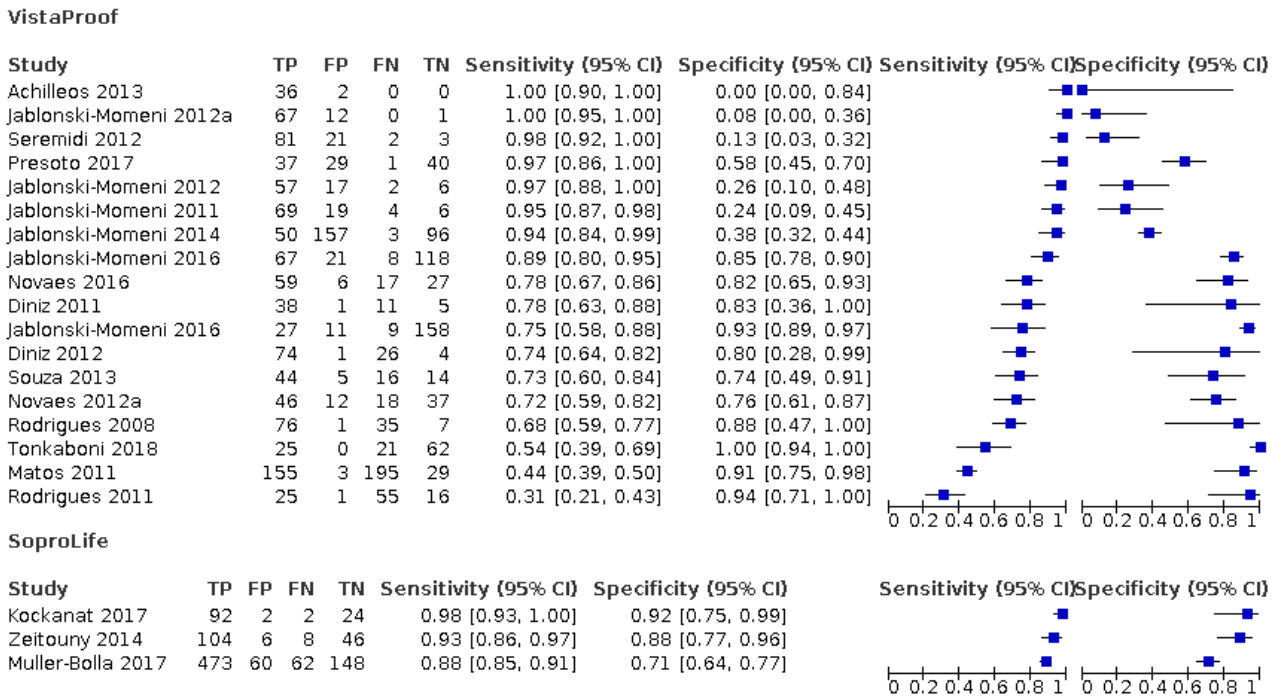
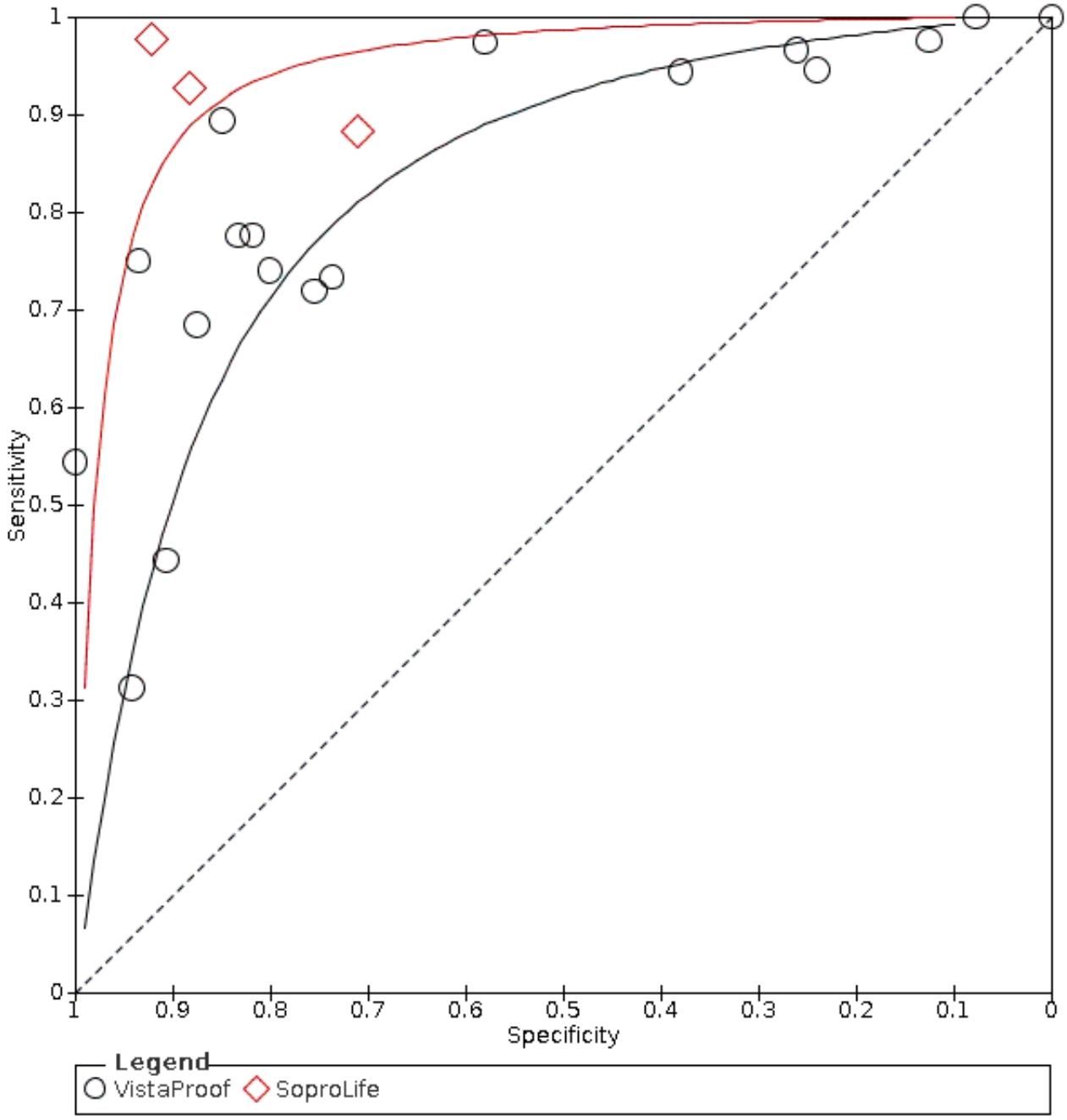


Figure 25. Summary receiver operating characteristic (SROC) plot of fluorescence tests investigating differences between red fluorescence devices VistaProof and SoproLife.



There was an observed difference between SoproLife and VistaProof at RDOR 4.75 (95% CI 1.46 to 15.45; P = 0.0095), however with only three included studies for SoproLife this result should be interpreted with caution.

Green fluorescence

We found nine studies that used green fluorescence methods (QLF) to detect caries (Bussaneli 2015; Diniz 2019; Feng 2005; Jung 2018; Kim 2017; Ko 2015; Lee 2018; Pereira 2011; Yoon 2017).The coupled forest plot is presented along with the estimates of sensitivity and specificity for each study and plotted in receiver operating characteristic (ROC) space (Figure 17; Figure 21). There was considerable variation in the estimates of both sensitivity and specificity, which covered the ranges 0.42 to 0.96 and 0.39 to 0.95 respectively. Individual study estimates of sensitivity and specificity are shown in Figure 26 and SROC estimates are shown in Figure 27. Two different approaches were apparent

in the QLF group, a software-based decision and an image-based decision, there was no significant difference between the results of these two groups RDOR 3.10 (95% CI 0.38 to 25.07; P = 0.34).

Figure 26. Forest plot of fluorescence tests investigating differences between green fluorescence devices with image and function.

Study	TP	FP	FN	TN	QLF function	Sensitivity (95% CI)	Specificity (95% CI)
Pereira 2011	53	25	2	16	Image	0.96 [0.87, 1.00]	0.39 [0.24, 0.55]
Feng 2005	342	75	15	1300	Software	0.96 [0.93, 0.98]	0.95 [0.93, 0.96]
Yoon 2017	61	21	3	17	Software	0.95 [0.87, 0.99]	0.45 [0.29, 0.62]
Jung 2018	333	183	39	236	Software	0.90 [0.86, 0.92]	0.56 [0.51, 0.61]
Lee 2018	40	1	10	11	Software	0.80 [0.66, 0.90]	0.92 [0.62, 1.00]
Ko 2015	57	3	19	16	Software	0.75 [0.64, 0.84]	0.84 [0.60, 0.97]
Bussaneli 2015	46	5	20	23	Software	0.70 [0.57, 0.80]	0.82 [0.63, 0.94]
Diniz 2019	45	4	21	18	Software	0.68 [0.56, 0.79]	0.82 [0.60, 0.95]
Kim 2017	72	16	98	94	Image	0.42 [0.35, 0.50]	0.85 [0.77, 0.91]

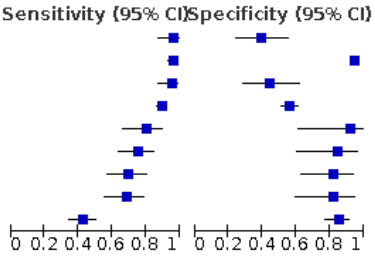
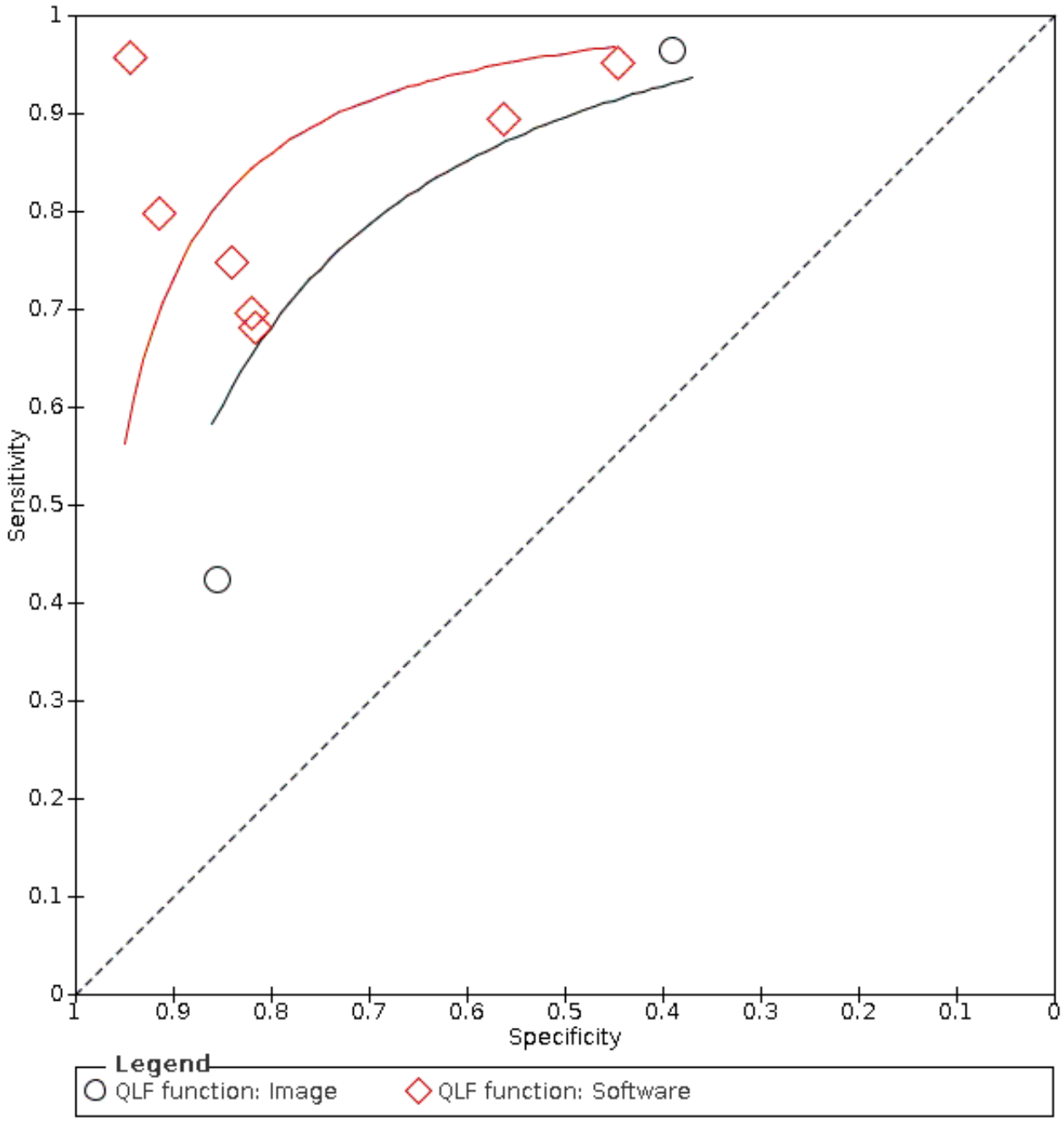


Figure 27. Summary receiver operating characteristic (SROC) plot of fluorescence tests investigating differences between green fluorescence devices with image and function.



HISTORY

Review first published: Issue 12, 2020

CONTRIBUTIONS OF AUTHORS

All review authors collaborated in the conception of the review purpose, design, and interpretation of results.
Drafting the protocol and final draft of the review: Tanya Walsh (TW), Richard Macey (RM).
Developing the search strategy: TW, RM.
Co-ordination of contributions from the co-authors: RM.

Screening of papers against eligibility criteria: RM, TW, Philip Riley (PR), Helen Worthington (HW), and Anne-Marie Glenny (AMG).
Obtained data on published, ongoing, and unpublished studies: RM.
Appraising the quality of papers: RM, TW, PR, HW, and AMG.
Extracting data for the review: RM, TW, PR, HW, Patrick Fee (PF), and AMG.
Entering data into Review Manager 5: RM.
Analysis of data: RM and TW.
Provided clinical guidance during all phases of review: Janet Clarkson (JC) and David Ricketts (DR).

DECLARATIONS OF INTEREST

Richard Macey: none known.
Tanya Walsh: none known. I am Statistical Editor with Cochrane Oral Health.
Philip Riley: none known. I am Deputy Co-ordinating Editor of Cochrane Oral Health.
Anne-Marie Glenny: none known. I am Co-ordinating Editor of Cochrane Oral Health.
Helen V Worthington: none known. I am an Editor with Cochrane Oral Health.
Patrick A Fee: none known.
Janet E Clarkson: none known. I am Co-ordinating Editor of Cochrane Oral Health.
David Ricketts: none known.

SOURCES OF SUPPORT

Internal sources

- Division of Dentistry, School of Medical Sciences, Faculty of Biology, Medicine and Health, The University of Manchester, UK
- Manchester Academic Health Sciences Centre (MAHSC) and the NIHR Manchester Biomedical Research Centre, UK

External sources

- National Institute for Health Research (NIHR), UK

  NIHR Cochrane Programme Grant: 16/114/23 Detection and diagnosis of common oral diseases: diagnostic test accuracy of tests of oral cancer and caries
- Cochrane Oral Health Global Alliance, Other

  The production of Cochrane Oral Health reviews has been supported financially by our Global Alliance since 2011 (oralhealth.cochrane.org/partnerships-alliances). Contributors in the last 2 years have been the American Association of Public Health Dentistry, USA; AS-Akademie, Germany; the British Association for the Study of Community Dentistry, UK; the British Society of Paediatric Dentistry, UK; the Canadian Dental Hygienists Association, Canada; the Centre for Dental Education and Research at All India Institute of Medical Sciences, India; the National Center for Dental Hygiene Research & Practice, USA; New York University College of Dentistry, USA; and Swiss Society of Endodontology, Switzerland
- NIHR, UK

  This project was supported by the NIHR, via Cochrane Infrastructure funding to Cochrane Oral Health. The views and opinions expressed herein are those of the review authors and do not necessarily reflect those of the Evidence Synthesis Programme, the NIHR, the NHS, or the Department of Health and Social Care

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

- Three categories of fluorescence index test were defined in the index test section of the 'Background'. It was important to categorise the devices as they each utilise different wavelengths to reach a diagnostic decision.
- One of the objectives was removed because the search produced a large body of evidence for the primary time point in clinical process so we decided it would add unnecessary complexity to investigate the additional objective of the value of each index test at different positions in the clinical pathway.
- We removed the secondary objective which stipulated that we would investigate the impact of previously applied restorations and fissure sealants as there were insufficient studies that included previously restored or sealed teeth. This also allowed us to amend the listed target conditions which stated caries adjacent to existing restorations.
- The protocol specified that we would investigate the difference between in vitro and in vivo studies, this has not been reported explicitly because the reference standard investigations cover the same issue. All in vitro studies employed a histological reference standard so this can be used as a proxy for the in vitro/in vivo comparison.